

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2004, 08:50:19 ; Search time 2133 Seconds

(without alignments)
10160.111 Million cell updates/sec

Title: US-09-720-934-1_COPY_1_500

Perfect score: 500
Sequence: 1 caaagaattccgggtacg9.....tatacgtaccctctgcact 500

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 3470272 segs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seg length: 0
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```
GenEmbl: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vt: *
15: em_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vt: *
30: em_hcg_hum: *
31: em_hcg_inv: *
32: em_hcg_other: *
33: em_hcg_mus: *
34: em_hcg_pln: *
35: em_hcg_rtd: *
36: em_hcg_mam: *
37: em_hcg_vrt: *
38: em_sv: *
39: em_hlgo_hum: *
40: em_hlgo_mus: *
41: em_hlgo_other: *
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	500	100.0	5199	6	BD205033 Isolated
2	486	97.2	877	6	BD124639 Primer fo
3	486	97.2	877	6	BD126208 Primer fo
4	486	97.2	1299	9	BC039036 Homo sapi
5	486	97.2	1676	6	BD127640 Primer fo
6	486	97.2	1676	9	AK075290 Homo sapi
7	486	97.2	2131	6	AX880753 Sequence
8	486	97.2	2131	6	BD127168 Primer fo
9	486	97.2	2131	6	BD158570 Primer fo
10	486	97.2	2131	9	AK027846 Homo sapi
11	486	97.2	2131	9	AK074554 Homo sapi
12	486	97.2	2199	9	BC058925 Homo sapi
13	486	97.2	5195	6	BD205035 Isolated
14	486	97.2	5381	9	AF114488 Homo sapi
15	486	97.2	5458	6	BD205034 Isolated
16	486	97.2	6439	9	AF114487 Homo sapi
17	471.4	94.3	676	6	AX870076 Sequence
18	471.4	94.3	676	6	BD150138 Primer fo
19	428	85.6	2079	6	BD205036 Isolated
20	399	79.8	5287	9	AF064243 Homo sapi
21	399	79.8	7247	9	AF064244 Homo sapi
22	280.2	56.0	3812	10	AF132672 Rattus no
23	280.2	56.0	4025	10	AF127798 Rattus no
24	241.8	48.4	3723	10	AF132478 Mus muscu
25	241.8	48.4	5145	10	AF132481 Mus muscu
26	239.8	48.0	493	10	AF525079 Mus muscu
27	216.6	43.3	4103	5	AF032118 Xenopus l
28	206	41.2	3390	9	HSM809003 Homo sapi
29	183	36.6	481	10	AF468654 Mus muscu
30	163.2	32.6	37425	9	AP000308 Homo sapi
31	163.2	32.6	57688	9	AP000048 Homo sapi
32	163.2	32.6	100000	9	AP000115 Homo sapi
33	163.2	32.6	100000	9	AP000191 Homo sapi
34	163.2	32.6	340000	9	AP001717 Homo sapi
35	161.6	32.3	163049	9	BS000182 Pan trogl
36	161.6	32.3	199436	9	BS000181 Pan trogl
37	156.8	31.4	157138	2	AC144366 Papio anu
38	154.4	30.9	1749	9	BC020921 Homo sapi
39	154.4	30.9	2085	9	BC038963 Homo sapi
40	154.4	30.9	4557	9	AF182199 Homo sapi
41	154.4	30.9	5938	9	AB033082 Homo sapi
42	154.4	30.9	6091	9	AF182198 Homo sapi
43	152.8	30.6	5828	6	BD167848 Method fo
44	152.8	30.6	5828	9	AF248540 Homo sapi
45	137.8	27.6	531	6	BD264620 Compositi

ALIGNMENTS

```
RESULT 1
BD205033
LOCUS      BD205033
DEFINITION Isolated SH3 gene relating to myeloproliferative disorders and
            leukemia and utilization thereof.
ACCESSION  BD205033
VERSION    BD205033.1 GI:33014803
KEYWORDS   JP 2002511267-A/1.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 5199)
AUTHORS   Korenberg, J.R. and Chen, X.N.
TITLE     Isolated SH3 gene relating to myeloproliferative disorders and
```

JOURNAL Patient: JP 2002511267-A/1.16-APR-2002;
CEDARS SINAI HEALTH SYSTEM ET AL
COMMENT OS Homo sapiens (human)
PN JP 2002511267-A/1
PD 16-APR-2002
PF 16-APR-1999 JP 2000543610
PR 16-APR-1998 US 60/0882007
PI JUIE R KORENBERG, XIAO NING CHEN
PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,
PC C12Q1/68,
PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and

leukemia
CC and utilization thereof.
FH Key Location/Qualifiers
FT source 1..5199
FT Location/Qualifiers
1..5199
/organism='Homo sapiens (human)'.
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN
Query Match 100.0%; Score 500; DB 6; Length 5199;
Best Local Similarity 100.0%; Pred. No.3.9e-120; Indels 0; Gaps 0;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

FEATURES
source
1 CAAAGAAATTCGCGGTACGCGCGCGCGAGAGAAATCCCGAGCGGCTCCCGGACGGA 60
1 CAAAAGAAATTCGCGGTACGCGCGCGCGCGAGAGAAATCCCGAGCGGCTCCCGGACGGA 60
61 CAGAGAGCGGCGCGGATGCTGTGCGGCGCTCGGCTCTCTCCCGACGCGCGCG 120
61 CAGAGAGCGGCGGCGGATGCTGTGCGGCGCTCGGCTCTCTCCCGACGCGCGCG 120
121 GAGAGGCGGCGGATGCTGTGCGGCGCGCGGCGGCGCGCGCGGAGATGAGCGCTC 180
121 GAGAGGCGGCGGATGCTGTGCGGCGCGCGGCGGCGCGCGCGGAGATGAGCGCTC 180
121 GAGAGGCGGCGGATGCTGTGCGGCGCGCGGCGGCGCGCGCGGAGATGAGCGCTC 180
181 GATTACCAAGGTAAAGTAAACAGAACCATGCTCAGTTTCCAAACCTTTTGTGTCAGC 240
181 GATTACCAAGGTAAAGTAAACAGAACCATGCTCAGTTTCCAAACCTTTTGTGTCAGC 240
241 CTGGATATCTGGGCGCTAACTGTAGAGAGAAAGCGGAGATGATGACAGTTCCATAGT 300
241 CTGGATATCTGGGCGCTAACTGTAGAGAGAAAGCGGAGATGATGACAGTTCCATAGT 300
301 TTAAGGCAATATCTGGATTCATTACTGGGATCAAGCTAGAAACTTTTTTTCATCT 360
301 TTAAGGCAATATCTGGATTCATTACTGGGATCAAGCTAGAAACTTTTTTTCATCT 360
361 GGGTTACCTCAACCTGTTTTTSCACAGATATGGGCACTAGCTGACATGAATATGATGA 420
361 GGGTTACCTCAACCTGTTTTTSCACAGATATGGGCACTAGCTGACATGAATATGATGA 420
421 AGAATGATCAATGAGATTTCCATAGCTATGAATCTATCAACTGAAGCTACAAAGA 480
421 AGAATGATCAATGAGATTTCCATAGCTATGAATCTATCAACTGAAGCTACAAAGA 480
481 TATCAGTACCTCTGCACT 500
481 TATCAGTACCTCTGCACT 500

RESULT 2
LOCUS BD124639 877 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD124639
VERSION BD124639.1 GI:23219584

KEYWORDS JP 2002017375-A/70.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS
1 (bases 1 to 877)
Oca,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE
JOURNAL
Patent: JP 2002017375-A 70 22-JAN-2002;
HELIIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002017375-A/70
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OCA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI ISHII,
PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI,HISASHI KOGA

FEATURES
source
1..877
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN
Query Match 97.2%; Score 486; DB 6; Length 877;
Best Local Similarity 100.0%; Pred. No.1.4e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

FEATURES
source
15 GTACGGCGGCTCGCGGAGAAATCCCGAGCGGCTCCCGGAGCGACAGAGAGCGGCGG 74
12 GTACGGCGGCTCGCGGAGAAATCCCGAGCGGCTCCCGGAGCGACAGAGAGCGGCGG 71
75 GGGATGATGCGGGGCTGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 134
72 GGGATGATGCGGGGCTGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 131
135 TTGTCCCTGGGCGGCGGAGCGGAGCCCGCGGAGATGAGGCGTCAATAGCAAGTAA 194
132 TTGTCCCTGGGCGGCGGAGCGGAGCCCGCGGAGATGAGGCGTCAATAGCAAGTAA 191
195 AAGTACAGAACCATGCTCAGTTTCCAAACCTTTTGTGAGAGCTGATATCTGGGC 254
192 AAGTACAGAACCATGCTCAGTTTCCAAACCTTTTGTGAGAGCTGATATCTGGGC 251
255 CATTAAGTGAAGAGAGAGAGAGATGATGACAGCTTCATAGTTTAAAGCAATATC 314
252 CATTAAGTGAAGAGAGAGAGAGATGATGACAGCTTCATAGTTTAAAGCAATATC 311
315 TGGATTCATTACTGGTATCAAGCTAGAAACTTTTTTTCATCTGGGTTACTCAAC 374
312 TGGATTCATTACTGGTATCAAGCTAGAAACTTTTTTTCATCTGGGTTACTCAAC 371
375 TGGTTAGACAGATATGGGCACTAGCTGACATGAATATGAGAGATGAGTCAAGT 434
372 TGGTTAGACAGATATGGGCACTAGCTGACATGAATATGAGAGATGAGTCAAGT 431
435 GGAATTTCCATAGTATGAATCTTAAACTGAAGCTACAGAGATATGACTACCTC 494
432 GGAATTTCCATAGTATGAATCTTAAACTGAAGCTACAGAGATATGACTACCTC 491
495 TGCCT 500


```

Db      492  TGCAC  497
|||||
RESULT 3
LOCUS   BD126208                877 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD126208.1 GI:23221153
VERSION   BD126208.1
KEYWORDS  JP 2002017375-A/1639.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS  Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
           Makamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
           Koga,H.
TITLE     Primer for synthesizing full-length cDNA and use thereof
JOURNAL   Patent: JP 2002017375-A 1639 22-JAN-2002;
           HELIX RESEARCH INSTITUTE
COMMENT   OS Homo sapiens (human)
           PN JP 2002017375-A/1639
           PD 22-JAN-2002
           PF 07-JUL-2000 JP 2000253172
           PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
           PI ISHII,
           PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
           SHINICHI KOJIMA,
           PI TETSUJI OTSUKI,HISASHI KOGA
           PC
           C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
           10,
           PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
           Primer for synthesizing full-length cDNA and use thereof FH Key
           Location/Qualifiers
           FT source 1..877
           /organism="Homo sapiens (human)"
           /location="Qualifiers"
           /mol_type="genomic DNA"
           /db_xref="taxon:9606"
ORIGIN
Query Match 97.2%; Score 486; DB 6; Length 877;
Best Local Similarity 100.0%; Pred. No. 1.4e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
15 GTACGGCGGCTCGCGAGAGAGATCCGAGCGGCTCGCGAGCGAGACGAGAGCGGCG 74
12 GTACGGCGGCTCGCGAGAGAGATCCGAGCGGCTCGCGAGCGAGACGAGAGCGGCG 71
75 GGGATGGGTGCGGGGCTCGCGGCTCTGCTCCCTCCAGCGGCGCGTGAAGGGCAGTGA 134
72 GGGATGGGTGCGGGGCTCGCGGCTCTGCTCCCTCCAGCGGCGCGTGAAGGGCAGTGA 131
135 TTGTGCTCGGGGCGGCGAGCGCGGACCCGCCGAGATGAGAGCGTGAATTAGCAAGTGA 194
132 TTGTGCTCGGGGCGGCGAGCGCGGACCCGCCGAGATGAGAGCGCGTGAATTAGCAAGTGA 191
195 AAGTAAACGAAACCATGCTCATGTTCCAAACCTTTGGTGGAGCGCTGAGATATTTGGGC 254
192 AAGTAAACGAAACCATGCTCATGTTCCAAACCTTTGGTGGAGCGCTGAGATATTTGGGC 251
255 CATAACTGTAGAGGAAAGAGGAGCATGATGACAGTTCCATAGTTTAAAGCCAAATTC 314
252 CATAACTGTAGAGGAAAGAGGAGCATGATGACAGTTCCATAGTTTAAAGCCAAATTC 311
315 TGGATTCATTACTGTGATCAAGCTAGAACTTTTTCATCTGGGTTACCTCAACC 374
312 TGGATTCATTACTGTGATCAAGCTAGAACTTTTTCATCTGGGTTACCTCAACC 371

```

```

OY      375  TGTTCACACAGATATGGGACACTGACATGATATATGATGAGAAATGATCAACT 434
Db      372  TGTTCACACAGATATGGGACACTGACATGATATATGATGAGAAATGATCAACT 431
OY      435  GGAGTTTCCATGACTATGAACCTTATCAACTGAGCTACAGAGATATCGTACCCCTC 494
Db      432  GGAGTTTCCATGACTATGAACCTTATCAACTGAGCTACAGAGATATCGTACCCCTC 491
OY      495  TGCAC  500
Db      492  TGCAC  497
|||||
RESULT 4
LOCUS   BC039036                1299 bp    mRNA     linear   PRI 24-DEC-2002
DEFINITION Homo sapiens, similar to intersectin 1 (SH3 domain protein), clone
ACCESSION BC039036
VERSION   BC039036.1 GI:24659505
KEYWORDS  partial cds.
SOURCE    BC039036.1
ORGANISM  Homo sapiens (human)
           Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS  Strausberg,R.
TITLE     Direct Submission
JOURNAL   Submitted (01-NOV-2002) National Institutes of Health, Mammalian
           Gene Collection (MGC), Cancer Genomics Office, National Cancer
           Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
           USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
           Email: cga@bbs-remail.nih.gov
           Tissue Procurement: ATCC
           cDNA library Preparation: Life Technologies, Inc.
           cDNA library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
           DNA Sequencing by: National Institutes of Health Intramural
           Sequencing Center (NISC),
           Gaithersburg, Maryland;
           Web site: http://www.nisc.nih.gov/
           Contact: nisc_mgc@nhi.nih.gov
           Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
           Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
           Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
           Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R.,
           Maduro,Q.L., Mastello,C., Maekel,B., Mastrian,S.D., McCloskey,J.C.,
           McDowell,J., Pearson,R., Stantirpop,S., Thomas,P.J., Touchman,J.W.,
           Tsurgoun,C., Vogt,U.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
           Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNI at: http://image.llnl.gov
Series: IRAP Plate: 82 Row: 1 Column: 2
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4504796.
FEATURES
source 1..1299
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5784342"
/tissue_type="uterus", "leiomyosarcoma"
/clone_id="N1H MGC_71"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
264..>1299
/codon_start=1
/product="Similar to intersectin 1 (SH3 domain protein)"
/protein_id="AAH39036.1"
/db_xref="GI:27371135"
CDS

```

/translation="AAOPPTPPGSLDIWALTVEREAKRDOOHSLKPISGFITGDA
RMFPQSGAPPOPVLAQIMALAMNNDGMDVEFSIAMLIKIKIAGYOLPALPBPV
KQPPVIASSAPAFKAGIAMSPLUTVAVAPVMSITPVYKMSPTLUSVTTAAVPLAN
GAPVYQPLPAPAHPPATLTKSSSFSRSGPQSLNKLKQASFDVASPVAEMVAV
OSSRLKRYQLNSHDKTMSGHLGPQARTILMOSLPQQLASIMLSIDIDDKLTA
EEFIILAMHILIDVAMSQPLPVPVLPPEYIPPSFRVRSVSGISVISTSDORLPEPV
LDEDOOKKKK"

ORIGIN

Query Match 97.2%; Score 486; DB 9; Length 1299;
Best Local Similarity 100.0%; Pred. No. 1.5e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	15	GTACGGCGGCTGCGGAGAAAGATCCCGAGCGGGCTCCGGGACGACAGAGAGCGGGCG	74
DB	71	GTACGGCGGCTGCGGAGAAAGATCCCGAGCGGGCTCCGGGACGACAGAGAGCGGGCG	130
QY	75	GGGATGCTGTGCGGGGCTGCGGCTCCGCTCCGCTCCGAGCGGCGGTGAGCGGCACTGA	134
DB	131	GGGATGCTGTGCGGGGCTGCGGCTCCGCTCCGCTCCGAGCGGCGGTGAGCGGCACTGA	190
QY	135	TTTGTCCCTGGGGCGGCGAGCGGTGACCCCGCGAGATGAGCGGTGATTAGCAAGTAA	194
DB	191	TTTGTCCCTGGGGCGGCGAGCGGTGACCCCGCGAGATGAGCGGTGATTAGCAAGTAA	250
QY	195	AACTAACAGAACCATGCTCACTTTCCACACCTTTTGGTGGCGAGCGCTGATATCTGGGG	254
DB	251	AACTAACAGAACCATGCTCACTTTCCACACCTTTTGGTGGCGAGCGCTGATATCTGGGG	310
QY	255	CATTAAGTGAAGAAAGCGGAGAGATGATGAGCGGTTCATAGTTTAAAGCAATATC	314
DB	311	CATTAAGTGAAGAAAGCGGAGAGATGATGAGCGGTTCATAGTTTAAAGCAATATC	370
QY	315	TGATTCATTACCTGCGATCAAGCTAGAAAATTTTTTTCATCTGGGTACTCAAC	374
DB	371	TGATTCATTACCTGCGATCAAGCTAGAAAATTTTTTTCATCTGGGTACTCAAC	430
QY	375	TGTTTAAAGCAGATATGAGCACTAGCTGACATGAATATGAGAAAGATGATCAAGT	434
DB	431	TGTTTAAAGCAGATATGAGCACTAGCTGACATGAATATGAGAAAGATGATCAAGT	490
QY	435	GGAGTTTCCATAGCTATGAAACTTATCAAACTGAAGCTACAGAGATATGATCACTC	494
DB	491	GGAGTTTCCATAGCTATGAAACTTATCAAACTGAAGCTACAGAGATATGATCACTC	550
QY	495	TGCACT 500	
DB	551	TGCACT 556	

RESULT 5
BD127640 1676 bp DNA linear PAT 18-SEP-2002

LOCUS BD127640
DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD127640
VERSION BD127640.1 GI:23222585
KEYWORDS JP 2002017375-A/3071.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1676)
Ota, T., Nishikawa, T., Isegai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Makamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.

TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 3071 22-JAN-2002;
HELIIX RESEARCH INSTITUTE

COMMENT OS Homo sapiens (human)
PN JP 2002017375-A/3071
PD 22-JAN-2002 JP 2002253172
PF 07-JUL-2000 JP 2002253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISEGAI, KOJI HAYASHI, SHIZUKO

PI ISHII,
PI YURI KAWAI, AI MAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOIUMA,
PI TETSUO OTSUKI, HISASHI KOGA
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
(264). (1676).

FEATURES

source 1..1676
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 97.2%; Score 486; DB 6; Length 1676;
Best Local Similarity 100.0%; Pred. No. 1.5e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	15	GTACGGCGGCTGCGGAGAAAGATCCCGAGCGGGCTCCGGGACGACAGAGAGCGGGCG	74
DB	71	GTACGGCGGCTGCGGAGAAAGATCCCGAGCGGGCTCCGGGACGACAGAGAGCGGGCG	130
QY	75	GGGATGCTGTGCGGGGCTGCGGCTCCGCTCCGCTCCGAGCGGCGGTGAGCGGCACTGA	134
DB	131	GGGATGCTGTGCGGGGCTGCGGCTCCGCTCCGCTCCGAGCGGCGGTGAGCGGCACTGA	190
QY	135	TTTGTCCCTGGGGCGGCGAGCGGTGACCCCGCGAGATGAGCGGTGATTAGCAAGTAA	194
DB	191	TTTGTCCCTGGGGCGGCGAGCGGTGACCCCGCGAGATGAGCGGTGATTAGCAAGTAA	250
QY	195	AACTAACAGAACCATGCTCACTTTCCACACCTTTTGGTGGCGAGCGCTGATATCTGGGG	254
DB	251	AACTAACAGAACCATGCTCACTTTCCACACCTTTTGGTGGCGAGCGCTGATATCTGGGG	310
QY	255	CATTAAGTGAAGAAAGCGGAGAGATGATGAGCGGTTCATAGTTTAAAGCAATATC	314
DB	311	CATTAAGTGAAGAAAGCGGAGAGATGATGAGCGGTTCATAGTTTAAAGCAATATC	370
QY	315	TGATTCATTACCTGCGATCAAGCTAGAAAATTTTTTTCATCTGGGTACTCAAC	374
DB	371	TGATTCATTACCTGCGATCAAGCTAGAAAATTTTTTTCATCTGGGTACTCAAC	430
QY	375	TGTTTAAAGCAGATATGAGCACTAGCTGACATGAATATGAGAAAGATGATCAAGT	434
DB	431	TGTTTAAAGCAGATATGAGCACTAGCTGACATGAATATGAGAAAGATGATCAAGT	490
QY	435	GGAGTTTCCATAGCTATGAAACTTATCAAACTGAAGCTACAGAGATATGATCACTC	494
DB	491	GGAGTTTCCATAGCTATGAAACTTATCAAACTGAAGCTACAGAGATATGATCACTC	550
QY	495	TGCACT 500	
DB	551	TGCACT 556	

RESULT 6
AK075290 1676 bp mRNA linear PRI 03-SEP-2002

LOCUS AK075290
DEFINITION Homo sapiens cDNA FL90809 f1s, clone Y79AA1000778, highly similar
to Homo sapiens interseccin long form mRNA.

ACCESSION AK075290
VERSION AK075290.1 GI:22761283
KEYWORDS Oligo capping; f1s (full insert sequence).

SOURCE OS Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
Isegai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T.,


```

Db      574 GGAATTTCCATACCTATGAAACTTATCAACTAGAGTACAGAGATATACGTAACCTC 633
Qy      495 TGCACT 500
Db      634 TGCACT 639

RESULT 8
BD127168
LOCUS   BD127168
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127168
VERSION   BD127168.1 GI:23222113
KEYWORDS  JP 2002017375-A/2599.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
          1 (bases 1 to 2131)
          Oca,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
          Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
          Koga,H.
          Primer for synthesizing full-length cDNA and use thereof
          Patent: JP 2002017375-A 2599 22-JAN-2002;
          HELIX RESEARCH INSTITUTE
          OS Homo sapiens (human)
          PN JP 2002017375-A/2599
          PD 22-JUN-2002
          PF 07-JUL-2000 JP 20020253172
          PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
          PI ISHII,
          PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
          SHINICHI KOJIMA,
          PI TETSUOI OTSUKI,HISASHI KOGA
          PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/
          10,
          PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
          Primer for synthesizing full-length cDNA and use thereof FH key
          Location/Qualifiers
          FT CDS
          source      1..2131
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

ORIGIN
Query Match      97.2%; Score 486; DB 6; Length 2131;
Best Local Similarity 100.0%; Pred. No. 1.6e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      15 GTACGGCGGCTCGCGAGAAAGAAATCCCGAGCGGCTCCCGGACGAGACAGAGCGGGCG 74
Db      12 GTACGGCGGCTCGCGAGAAAGAAATCCCGAGCGGCTCCCGGACGAGACAGAGCGGGCG 71
Qy      75 GGGATGATGTCGGGGGCTGCGGCTCTGCTCTCCCTCCACGCGGCGGTGAGCGGCACTGA 134
Db      72 GGGATGATGTCGGGGGCTGCGGCTCTGCTCTCCCTCCACGCGGCGGTGAGCGGCACTGA 131
Qy      135 TTTGTCCTTGCGGGGCGGCGCGCGGACCCGCGCGAGATGAGCGGTGAGCAAGGTAA 194
Db      132 TTTGTCCTTGCGGGGCGGCGCGCGGACCCGCGCGAGATGAGCGGTGAGCAAGGTAA 191
Qy      195 AAGTAACAGAACCATGCTCACTTTTCCAAACCTTTTGGTGGACGCTGATATCTGGGC 254
Db      192 AAGTAACAGAACCATGCTCACTTTTCCAAACCTTTTGGTGGACGCTGATATCTGGGC 251
Qy      255 CATAACTGTAGAGGAAGAAGCGCAAGCATGATGAGAGTTCCATATGTTTAAAGCAATATC 314
Db      252 CATAACTGTAGAGGAAGAAGCGCAAGCATGATGAGAGTTCCATATGTTTAAAGCAATATC 311

```

```

Qy      315 TGATTCATTACTGTGATGATCAAGTAGAACTTTTTTCAATCGGTACTCAACC 374
Db      312 TGATTCATTACTGTGATGATCAAGTAGAACTTTTTTCAATCGGTACTCAACC 371
Qy      375 TGTTTAGACACAGATATGGGCACTAGCTGACATGATATGATGAGAAATGATCAAGT 434
Db      372 TGTTTAGACACAGATATGGGCACTAGCTGACATGATATGATGAGAAATGATCAAGT 431
Qy      435 GGAATTTCCATACCTATGAAACTTATCAACTGAAAGCTACAGAGATATCACTACCTC 494
Db      432 GGAATTTCCATACCTATGAAACTTATCAACTGAAAGCTACAGAGATATCACTACCTC 491
Qy      495 TGCACT 500
Db      492 TGCACT 497

RESULT 9
BD158570
LOCUS   BD158570
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD158570
VERSION   BD158570.1 GI:27864328
KEYWORDS  JP 2002191363-A/13413.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
          1 (bases 1 to 2131)
          Oca,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,T.,
          Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
          Primer for synthesizing full-length cDNA and use thereof
          Patent: JP 2002191363-A 13413 09-JUL-2002;
          HELIX RESEARCH INSTITUTE
          OS Homo sapiens (human)
          PN JP 2002191363-A/13413
          PD 09-JUL-2002
          PF 28-JUL-2000 JP 2000280990
          PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
          PI SAITO,
          PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
          PI KEIICHI NAGAI,TETSUOI OTSUKI
          PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/
          10,
          PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
          Primer for synthesizing full-length cDNA and use thereof FH key
          Location/Qualifiers
          FT CDS
          source      1..2131
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

ORIGIN
Query Match      97.2%; Score 486; DB 6; Length 2131;
Best Local Similarity 100.0%; Pred. No. 1.6e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      15 GTACGGCGGCTCGCGAGAAAGAAATCCCGAGCGGCTCCCGGACGAGACAGAGCGGGCG 74
Db      154 GTACGGCGGCTCGCGAGAAAGAAATCCCGAGCGGCTCCCGGACGAGACAGAGCGGGCG 213
Qy      75 GGGATGATGTCGGGGGCTGCGGCTCTGCTCTCCCTCCACGCGGCGGTGAGCGGCACTGA 134
Db      214 GGGATGATGTCGGGGGCTGCGGCTCTGCTCTCCCTCCACGCGGCGGTGAGCGGCACTGA 273
Qy      135 TTTGTCCTTGCGGGGCGGCGAGCGCGCGCGAGATGAGCGGTGAGCAAGGTAA 194
Db      274 TTTGTCCTTGCGGGGCGGCGAGCGCGCGCGAGATGAGCGGTGAGCAAGGTAA 333
Qy      195 AAGTAACAGAACCATGCTCACTTTTCCAAACCTTTTGGTGGACGCTGATATCTGGGC 254

```

```

Db      334  AAGTACAGAACCATGCTCAAGTTTCCAAACCTTTTGGTGGCAGCTGGATATCTGGGC 393
QY      255  CATACCTGTAGAGGAAAGAGCGAAGCATGATCAGAGCTTCATGTTTAAAGCCAAATATC 314
Db      334  CATACCTGTAGAGGAAAGAGCGAAGCATGATCAGAGCTTCATGTTTAAAGCCAAATATC 453
QY      315  TGGATTCTATTACTGGTGTACAGCTAGAAACCTTTTTCATCTGGGTTACCTCAACC 374
Db      454  TGGATTCTATTACTGGTGTATCAAGCTAGAAACCTTTTTCATCTGGGTTACCTCAACC 513
QY      375  TGGTTTACACAGATATGGGCGCTAGCTGACATGATATATATATGATGAGAGATGATCAAGT 434
Db      514  TGTTTTACACAGATATGGGCGCTAGCTGACATGATATATATGATGAGAGATGATCAAGT 573
QY      435  GGAGTTTTCATAGCTATGAAACTTATCAAACTGAGCTACAGAGATATCAAGTACCTC 494
Db      574  GGAGTTTTCATAGCTATGAAACTTATCAAACTGAGCTACAGAGATATCAAGTACCTC 633
QY      495  TGCACCT 500
Db      634  TGCACCT 639

RESULT 10
LOCUS   AK027846                2131 bp    mRNA    linear    PRI 01-AUG-2002
DEFINITION Homo sapiens cDNA FLJ14940 fis, clone PLACE1010942, highly similar
            to Homo sapiens intersectin long isoform (ITSN) mRNA.
ACCESSION AK027846
KEYWORDS  oligo capping; fis (full insert sequence).
SOURCE    Homo sapiens
           Human
ORGANISM  Homo sapiens
           Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
           Eukaryota; Metazoa; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS   1 Isogai,T., Oca,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
           Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K.,
           Tanai,H., Kimata,M., Watanabe,M., Hiraoaka,S., Ishii,S., Kawai,Y.,
           Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahara,K.,
           Masuo,Y. and Kanehori,K.
           NEDO human cDNA sequencing project
TITLE      Unpublished
JOURNAL    2 (bases 1 to 2131)
REFERENCE  Isogai,T. and Otsuki,T.
AUTHORS   Direct Submission
JOURNAL   Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
           Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
           (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT    NEDO human cDNA sequencing project supported by Ministry of
           Economy, Trade and Industry of Japan; cDNA full insert sequencing;
           Research Association for Biotechnology; cDNA library construction;
           5'- & 3'-end one pass sequencing and clone selection; Helix
           Research Institute (supported by Japan Key Technology Center etc.)
           and Department of Virology, Institute of Medical Science,
           University of Tokyo.
FEATURES
  source
    1..2131
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="PLACE1010942"
    /tissue_type="Placenta"
    /clone_lib="PLACEL1"
    /note="Cloning vector: pME186FL3"

ORIGIN
Query Match 97.2%; Score 486; DB 9; Length 2131;
Best Local Similarity 100.0%; Pred. No. 1.6e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

15 GTACGGCGGCTCGCGAGGAAGATCCCGAGCGGGCTCCGGAGCGAGCAGAGAGCGGGCG 74

```

```

Db      154  GTACGGCGGCTCGCGAGGAAGATCCCGAGCGGGCTCCGGAGCGAGCAGAGAGCGGGCG 213
QY      75  GGGATGATGTGCGGGGCTGCGGCTCTCGGTCCTCCAGCGGCGGTGAGCGGCACTGA 134
Db      214  GGGATGATGTGCGGGGCTGCGGCTCTCGGTCCTCCAGCGGCGGTGAGCGGCACTGA 273
QY      135  TTTGTCCCTGGGGCGGCGAGCCCGGACCCCGGAGATGAGGCGTCATTTAGCAAGTGA 194
Db      274  TTTGTCCCTGGGGCGGCGAGCCCGGACCCCGGAGATGAGGCGTCATTTAGCAAGTGA 333
QY      195  AAGTACAGAACCATGCTCAAGTTTCCAAACCTTTTGGTGGCAGCTGGATATCTGGGC 254
Db      334  AAGTACAGAACCATGCTCAAGTTTCCAAACCTTTTGGTGGCAGCTGGATATCTGGGC 393
QY      255  CATACCTGTAGAGGAAAGAGCGAAGCATGATCAGAGCTTCATGTTTAAAGCCAAATATC 314
Db      394  CATACCTGTAGAGGAAAGAGCGAAGCATGATCAGAGCTTCATGTTTAAAGCCAAATATC 453
QY      315  TGGATTCTATTACTGGTGTACAGCTAGAAACCTTTTTCATCTGGGTTACCTCAACC 374
Db      454  TGGATTCTATTACTGGTGTATCAAGCTAGAAACCTTTTTCATCTGGGTTACCTCAACC 513
QY      375  TGGTTTACACAGATATGGGCGCTAGCTGACATGATATATATGATGAGAGATGATCAAGT 434
Db      514  TGTTTTACACAGATATGGGCGCTAGCTGACATGATATATGATGAGAGATGATCAAGT 573
QY      435  GGAGTTTTCATAGCTATGAAACTTATCAAACTGAGCTACAGAGATATCAAGTACCTC 494
Db      574  GGAGTTTTCATAGCTATGAAACTTATCAAACTGAGCTACAGAGATATCAAGTACCTC 633
QY      495  TGCACCT 500
Db      634  TGCACCT 639

RESULT 11
LOCUS   AK074554                2131 bp    mRNA    linear    PRI 03-SEP-2002
DEFINITION Homo sapiens cDNA FLJ90073 fis, clone HEMBA1004110, highly similar
            to Homo sapiens intersectin short form mRNA.
ACCESSION AK074554
KEYWORDS  oligo capping; fis (full insert sequence).
SOURCE    Homo sapiens
           Human
ORGANISM  Homo sapiens
           Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
           Eukaryota; Metazoa; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS   1 Isogai,T., Oca,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
           Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
           Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
           Kojima,S., Nagahara,K., Masuo,Y., Ono,T., Okano,K., Yoshikawa,Y.,
           Aotake,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
           Nimomiya,K.
           NEDO human cDNA sequencing project
TITLE      Unpublished
JOURNAL    2 (bases 1 to 2131)
REFERENCE  Isogai,T. and Otsuki,T.
AUTHORS   Direct Submission
JOURNAL   Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
           Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
           (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT    NEDO human cDNA sequencing project supported by Ministry of
           Economy, Trade and Industry of Japan; cDNA full insert sequencing;
           Research Association for Biotechnology; cDNA library construction;
           Institute of Medical Science, University of Tokyo, Laboratory of
           Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
           sequencing and clone selection; Helix Research Institute (supported
           by Japan Key Technology Center etc.).
FEATURES
  source
    1..2131
    /organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEMBA1004110"
/tissue_type="whole embryo, mainly head"
/clone_id="HEMBA1"
/dev_stage="embryo, 10 weeks"
/note="cloning vector: pME18SFL3"

```

```

Query Match      97.2%; Score 486; DB 9; Length 2131;
Best Local Similarity 100.0%; Pred. No. 1,6e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 15 GTACGGGGGCTCGGAGAGAAATCCGAGGGGCTCCGGAGACGAGACAGAGGGCGG 74
DB 12 GTACGGGGGCTCGGAGAGAAATCCGAGGGGCTCCGGAGACGAGACAGAGGGCGG 71
QY 75 GGGATGTGTGGCGGGGCTGCGCTCTCTGCTCCCTCCAGCGGGCGGTGAGCGGCACTGA 134
DB 72 GGGATGTGTGGCGGGGCTGCGCTCTCTGCTCCCTCCAGCGGGCGGTGAGCGGCACTGA 131
QY 135 TTGTCCCTGGGGGCGAGAGCGGCGGCGCGGAGATGAGGGGTGATTAAGCAAGTAA 194
DB 132 TTGTCCCTGGGGGCGAGAGCGGCGGCGGAGATGAGGGGTGATTAAGCAAGTAA 191
QY 195 AAGTAAAGAGACATGAGCTCAATTTTCCAGACCTTTTGTGGAGCGCTGATATCTGGGC 254
DB 192 AAGTAAAGAGACATGAGCTCAATTTTCCAGACCTTTTGTGGAGCGCTGATATCTGGGC 251
QY 255 CATTAAGTGTAGAGAAAGAGCGGAGAGATGAGAGTCAATGATTTAAAGCCATATATC 314
DB 252 CATTAAGTGTAGAGAAAGAGCGGAGAGATGAGAGTCAATGATTTAAAGCCATATATC 311
QY 315 TGAATTCATATCTGGGTATGAGTGAATCTTTTTCATCTGGGTACTTCAACC 374
DB 312 TGAATTCATATCTGGGTATGAGTGAATCTTTTTCATCTGGGTACTTCAACC 371
QY 375 TGTTTTGAAGACAGATATGGGCTAGCTGACATGATATATGAGAGATGATCAAGT 434
DB 372 TGTTTTGAAGACAGATATGGGCTAGCTGACATGATATATGAGAGATGATCAAGT 431
QY 435 GGAATTTTCATATGCTATGAAACTTATCAAACTGAGACGTAAGAGATATCACTACCTC 494
DB 432 GGAATTTTCATATGCTATGAAACTTATCAAACTGAGACGTAAGAGATATCACTACCTC 491
QY 495 TGCACCT 500
DB 492 TGCACCT 497

```

```

RESULT 12
BC058925
LOCUS      2199 bp      mRNA      linear      PRI 05-NOV-2003
DEFINITION Homo sapiens interseclin 1 (SH3 domain protein), mRNA (cDNA clone
IMAGE:4443129), partial cds.
ACCESSION BC058925
VERSION    BC058925.1 GI:37589134
KEYWORDS
SOURCE
ORGANISM  Homo sapiens (human)
            BakaYota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 2199)
REFERENCE
AUTHORS    Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,
            Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
            Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
            Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
            Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
            Stalcup, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
            Scheetz, T.E., Brownstein, M.J., Utsidi, T.B., Toshiyuki, S.,
            Carninci, P., Prange, C., Paha, S.S., Loquellano, N.A., Peters, G.J.,
            Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.U.,
            McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

```

```

TITLE
JOURNAL
MEDLINE
PUBMED
AUTHORS
TITLE
JOURNAL

```

```

REMARK
COMMENT

```

```

FEATURES
source
    Clone distribution: MGC clone distribution information can be found
    through the I.M.A.G.E. Consortium/ILN at: http://image.llnl.gov
    Series: IRAP Plate: 119 Row: b Column: 20
    This clone was selected for full length sequencing because it
    passed the following selection criteria: matched mRNA gi: 4504796.
    location/Qualifiers
        1..2199
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4443129"
            /tissue_type="Liver, adenocarcinoma"
            /clone_lib="NIH MGC_90"
            /lab_host="DH10B"
            /note="vector: PCMV-SPOrt6"
            1..>2199
                /gene="ITSN1"
                /note="synonyms: ITSN, SH3D1A, SH3P17"
                /db_xref="LOCUSID:6453"
                /db_xref="MTM:602442"
                256..>2199
                    /gene="ITSN1"
                    /codon_start=1
                    /product="ITSN1 protein"
                    /protein_id="AAH58925.1"
                    /db_xref="GI:37589135"
                    /db_xref="LOCUSID:6453"
                    /db_xref="MTM:602442"

```

```

gene
CDS

```

```

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huylk, S.W.,
Vallalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Rahcy, J., Helton, E., Kerteman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schwartz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.U. and Marr, M.A.,
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 2199)
Strausberg, R.
Direct Submission
Submitted (01-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNM)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reto Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Matheson, Candice McLeavy, Steven
Ness, Pawan Pandori, Anna-Liisa Prahru, Parvaneh Saedi, Jacqueline
Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Mariana Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

```

misc_feature
335..619
/gene="TTSN1"
/note="EH; Region: Eps15 homology domain"
/db_xref="CDD:smart0027"

misc_feature
935..1171
/gene="TTSN1"
/note="EH; Region: Eps15 homology domain"
/db_xref="CDD:smart0027"

ORIGIN

Query Match 97.2%; Score 486; DB 9; Length 2199;
Best Local Similarity 100.0%; Pred. No. 1.6e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GTACGGCGGCTCGCGAGAGAGATCCGAGCGGCTCCGGAGCGAGCAGAGAGCGGCG 74
DB 103 GTACGGCGGCTCGCGAGAGAGATCCGAGCGGCTCCGGAGCGAGCAGAGAGCGGCG 162

QY 75 GGATGAGTGTGGGGGCTCGCGGCTCTGCTCCCTCCGAGCGGCGCGTGAAGCGGCACTGA 134
DB 163 GGATGAGTGTGGGGGCTCGCGGCTCTGCTCCCTCCGAGCGGCGCGTGAAGCGGCACTGA 222

QY 135 TTGTCCCTGGGGGCGGCGCGGCGGAGATGAGGCGCTCGATTAGCAAGTTAA 194
DB 223 TTGTCCCTGGGGGCGGCGCGGCGGAGATGAGGCGCTCGATTAGCAAGTTAA 282

QY 195 AAGTACAGAACCATGCTCTGCTTCCAAACCTTTTGGTGGAGCGCTGATATCTGGGC 254
DB 283 AAGTACAGAACCATGCTCTGCTTCCAAACCTTTTGGTGGAGCGCTGATATCTGGGC 342

QY 255 CATACCTAGAGAGAGAGAGAGATGATGAGATTCATGTTTAAAGCCAAATTC 314
DB 343 CATACCTAGAGAGAGAGAGATGATGAGATTCATGTTTAAAGCCAAATTC 402

QY 315 TGGATTCTACTGCTGATCAAGCTAGAACTTTTTCATCTGGGTTACCTCAAC 374
DB 403 TGGATTCTACTGCTGATCAAGCTAGAACTTTTTCATCTGGGTTACCTCAAC 462

QY 375 TGTTTAGCAGATATGGGCACTAGCTGACATGATATATGATGAGAAATGATCAAGT 434
DB 463 TGTTTAGCAGATATGGGCACTAGCTGACATGATATATGATGAGAAATGATCAAGT 522

QY 435 GGAGTTTCCATAGCTATGAACCTATCAACTGAAGCTACAGGATATCAGTCACTC 494
DB 523 GGAGTTTCCATAGCTATGAACCTATCAACTGAAGCTACAGGATATCAGTCACTC 582

QY 495 TGCACCT 500
DB 583 TGCACCT 588

RESULT 13
BD205035 5195 bp DNA linear PAT 17-JUL-2003
LOCUS Isolated SH3 gene relating to myeloproliferative disorders and
DEFINITION leukemia and utilization thereof.
ACCESSION BD205035
VERSION BD205035.1 GI:33014805
KEYWORDS JP 2002511267-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 5195)
AUTHORS Korenberg, J.R. and Chen, X.N.
TITLE Isolated SH3 gene relating to myeloproliferative disorders and
leukemia and utilization thereof
JOURNAL Patent: JP 2002511267-A 3 16-APR-2002;

COMMENT
OS CEDARS SINAI HEALTH SYSTEM ET AL
PN Homo sapiens (human)
EN JP 2002511267-A/3
PD 16-APR-2002
PF 16-APR-1999 JP 2000543610
PR 16-APR-1998 US 60/082007
PI JULIE R. KORENBERG, XIMO NING CHEN
PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,
PC C12N13/68,
PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and
leukemia
and utilization thereof.
FH Key Location/Qualifiers
FT source 1..5195
FT /organism="Homo sapiens (human)"
FT /mol_type="genomic DNA"
FT /db_xref="taxon:9606"

ORIGIN

Query Match 97.2%; Score 486; DB 6; Length 5195;
Best Local Similarity 100.0%; Pred. No. 1.6e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GTACGGCGGCTCGCGAGAGAGATCCGAGCGGCTCCGGAGCGAGCAGAGAGCGGCG 74
DB 46 GTACGGCGGCTCGCGAGAGAGATCCGAGCGGCTCCGGAGCGAGCAGAGAGCGGCG 105

QY 75 GGATGAGTGTGGGGGCTCGCGGCTCTGCTCCCTCCGAGCGGCGCGTGAAGCGGCACTGA 134
DB 106 GGATGAGTGTGGGGGCTCGCGGCTCTGCTCCCTCCGAGCGGCGCGTGAAGCGGCACTGA 165

QY 135 TTGTCCCTGGGGGCGGCGCGGCGGAGATGAGGCGCTCGATTAGCAAGTTAA 194
DB 166 TTGTCCCTGGGGGCGGCGCGGCGGAGATGAGGCGCTCGATTAGCAAGTTAA 225

QY 195 AAGTACAGAACCATGCTCTGCTTCCAAACCTTTTGGTGGAGCGCTGATATCTGGGC 254
DB 226 AAGTACAGAACCATGCTCTGCTTCCAAACCTTTTGGTGGAGCGCTGATATCTGGGC 285

QY 255 CATACCTAGAGAGAGAGAGATGATGAGATTCATGTTTAAAGCCAAATTC 314
DB 286 CATACCTAGAGAGAGAGAGATGATGAGATTCATGTTTAAAGCCAAATTC 345

QY 315 TGGATTCTACTGCTGATCAAGCTAGAACTTTTTCATCTGGGTTACCTCAAC 374
DB 346 TGGATTCTACTGCTGATCAAGCTAGAACTTTTTCATCTGGGTTACCTCAAC 405

QY 375 TGTTTAGCAGATATGGGCACTAGCTGACATGATATATGATGAGAAATGATCAAGT 434
DB 406 TGTTTAGCAGATATGGGCACTAGCTGACATGATATATGATGAGAAATGATCAAGT 465

QY 435 GGAGTTTCCATAGCTATGAACCTATCAACTGAAGCTACAGGATATCAGTCACTC 494
DB 466 GGAGTTTCCATAGCTATGAACCTATCAACTGAAGCTACAGGATATCAGTCACTC 525

QY 495 TGCACCT 500
DB 526 TGCACCT 531

RESULT 14
AF114488 5381 bp mRNA linear PRI 16-JUL-2002
LOCUS AF114488
DEFINITION Homo sapiens intersectin short isoform (TTSN) mRNA, complete cds.
ACCESSION AF114488
VERSION AF114488.1 GI:4808824
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases 1 to 5381) Pucharcos,C., Fuentes,J.J., Casas,C., de la Luna,S., Alcantara,S., Alfonso,M.L., Soriano,E., Estivill,X. and Pritchard,M. Alu-splice cloning of human Intersectin (ITSN), a putative multivalent binding protein expressed in proliferating and differentiating neurons and overexpressed in Down syndrome
TITLE	Eur. J. Hum. Genet. 7 (6), 704-712 (1999)
JOURNAL	99415290
MEDLINE	10482960
PUBMED	2 (bases 1 to 5381) Pucharcos,C., Fuentes,J.J., Pritchard,M. and Estivill,X. Direct Submission Submitted (16-DIC-1998) Medical and Molecular Genetics Center, Cancer Research Institute, U/Hospital de Llo., Avia. Castelldefels km. 2,7, Barcelona 08907, Spain
AUTHORS	Location/Qualifiers
JOURNAL	1..5381
FEATURES	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="21" /map="21q22.1-q22.2" 1..5381 /gene="ITSN" 269..3931 /gene="ITSN"
SOURCE	/codon_start=1 /product="Intersectin short isoform" /protein_id="AAD29953.1" /db_xref="gi:4808825" /translation="MAQPPRPFGSLDIWATTVTERAKHDQFSLKPSIGFTGDQA RNFPGSLPOPVLAIOMALADMMNDGMDGVESIAMKKILKLGOGLPSALIPWM GAPVLAISAPAGMGSIASMPITLAVAPVMGSIPIVMGSPPTIASSNPFAVPLIAN GAPVIOTLPAPFAHPATLTPESSSFSSSGPSSOINTLKIQKASTDVASVPFAVAFAV QSSELYKRLPNSHDKTWSGHLTGPDAITLMQSSLPOAQIASITNSDIDQDEKTLPA EEFLIAHMLIDVAMSGQLPVPVLPPEYIIPSEFRVRSGSIGSTVSVDOLPEEBPV LEEDQDLLEKIPLYTFEDKRKENFERGULETEKROALLBOORKEORLIJAEABOE RKREEROBERKRLIELEKOLEKRELROEEREREIEREKALEBROLWEEM RNRRLOLLNRNKEDIEDIVYLKARKKTLEFLALNHKHOLEKDILDIRLTTQC EIESSTKSRELRIRAEITHLQQLOLESOOMLERLIPKOINDLQKYQONSIRHLSLV TKRALAKEARHLRDQDLDEVEKETRSKIQIDIFNNLUKEIRETHNKQLOKOKS MEARLKOKEGERKIIELEKQKEAQRGAQGRDOMELHVOQBEOHORPRKLHEBK KREBSVKXKQGEENKQEOADQIKRLPHHQEPAPKVOAPSTAEGPLTISAQENV KVYEVRAIYPFESSHDEITTOPEDIIVHWKGEMVEDSQTEGPMWLGELIKGTGTPFA NYAEKIPENEYPAVYKPVDTSTAPAKRLARETPAPLAVTSRSPITPNMWADPSST WPNSTNEKPETDNWDAAWAQPSLTVPBAQOLRQSAFTPATRTATSSSPVLOGEEXBY GLQAOALYPMRAKKNHLNFKNDVITYILEQDMWMMGEVQGGMFPKSVKVLGISG IRKTSMDSSGESPALSKRVASPAAPVVGEEFIAMYTESSEQDLTFYQSDVILI VTKKDGWMWTGTVGDKAVEPSNVYRLKDSGSTAGTKTSLGKKPEIAOVASYTAN GPELDITLAPQOLIIRKNPDGMMWGEIOARGKRROIQWPFANYKLSIPGTSKITPT EPPESTIALAACVITGMVDTYAQNDDIELAFNKQIINVLNKEDPDMMKGEVNGQVLF PSNVYKLTITMDPSQ"
ORIGIN	
Query Match	97.2%; Score 486; DB 9; Length 5381;
Best Local Similarity	100.0%; Pred. No. 1.8e-116;
Matches	486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Dy	15 GTACGGGGCGCTCCGAGAGAAGAACCCGAGAGGGGCTCCGGAGCGACAGAGAGCGGGCG 74
Dd	76 GTACGGCGGGCTCCGAGAGAGACAATCCGAGGGGCTCCGGAGCGACAGAGAGCGGGCG 135
Dy	75 GGAGTAGTGTCGGGGGTGCGGCTCTGTCGCTCCCTCCAGCGAGCGCGTGAAGCGCACTGA 134
Dd	136 GGAGTAGTGTCGGGGGTGCGGCTCTCTGCGTCTCCAGCGAGCGCGTGAAGCGCACTGA 195
Dy	135 TTGTGTCCTGGGGGTGGAGAGCGGACCCGCCCGAGATGAGGCTCATTTAGCAAAGTAA 194
Dd	196 TTTGTGTCCTGGGGGTGGAGAGCGGACCCGCCCGAGATGAGGCTCATTTAGCAAAGTAA 255

QY	195	AAGPAAACGAACCATGGCTGACGTTTCCAAACCTTTTGGTGGGACGCTGGATATCTGGGC	254
Db	256	AAGPAAACGAACCATGGCTGACGTTTCCAAACCTTTTGGTGGGACGCTGGATATCTGGGC	315
QY	255	CATAACTGTAGAGGAAAGAGCGAAGCATGATGATCAGACAGTTCATAGTTTAAAGCCAAATATC	314
Db	316	CATAACTGTAGAGGAAAGAGCGAAGCATGATGATCAGACAGTTCATAGTTTAAAGCCAAATATC	375
QY	315	TGGATTATATCTGGTGTCAAGCTAAGAACTTTTTCATCTGGGTTACTCAAC	374
Db	376	TGGATTATATCTGGTGTCAAGCTAAGAACTTTTTCATCTGGGTTACTCAAC	435
QY	375	TGTTTTCAGCAGATATGGGACCTAGCTGCATATATGATGGAAGATGCATCAAGT	434
Db	436	TGTTTTCAGCAGATATGGGACCTAGCTGCATATATGATGGAAGATGCATCAAGT	495
QY	435	GGAAGTTTCCATAGCTATGAAACTTATCAACTGAAGCTCAAGATATACGTAACCTC	494
Db	496	GGAAGTTTCCATAGCTATGAAACTTATCAACTGAAGCTCAAGATATACGTAACCTC	555
QY	495	TGGACT 500	
Db	556	TGGACT 561	
RESULT 15			
BD205034		5458 bp	DNA linear PAT 17-JUL-2003
LOCUS			
DEFINITION			Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.
ACCESSION			BD205034
VERSION			BD205034.1 GI:33014804
KEYWORDS			JP 2002511267-A/2.
SOURCE			Homo sapiens (human)
ORGANISM			Homo sapiens
REFERENCE			Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS			1 (bases 1 to 5458)
TITLE			Korenberg,J.R. and Chen,X.N.
JOURNAL			Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof
COMMENT			Patent: JP 2002511267-A 2 16-APR-2002; CEDARS SINAI HEALTH SYSTEM ET AL
			OS Homo sapiens (human)
			PN JP 2002511267-A/2
			PD 16-APR-2002
			PF 16-APR-1999 JP 2000543610
			PR 16-APR-1998 US 60/082007
			PT JULIE R KORENBERG,XIAO NING CHEN
			PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,
			PC C1201/68,
			PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
			CC Isolated SH3 gene relating to myeloproliferative disorders and
			leukemia
			CC and utilization thereof.
			FH Key Location/Qualifiers
			FT source 1..5458
			FT Location/Qualifiers
			1..5458 /organism='Homo sapiens (human)'. /mol_type='genomic DNA' /db_xref='taxon:9606'
FEATURES			
source			
ORIGIN			
Query Match			97.2%; Score 486; DB 6; Length 5458;
Best Local Similarity			100.0%; Pred. No. 1,8e-116;
Matches			486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	15	GTAACGCGAGCTCCAGAGAAATCCGACGGGCTCCGGACGGAACAAGACCGGGCG	74
Db	74	GTAACGCGAGCTCCAGAGAAATCCGACGGGCTCCGGACGGAACAAGACCGGGCG	133


```
QY 75 GGGATGCTGTGCGGGGCTGCGGCTCTGCGTCCCTCCAGCGCGCGGTGAGCGGCATCTGA 134
|||
Db 134 GGGATGCTGTGCGGGGCTGCGGCTCTGCGTCCCTCCAGCGCGCGGTGAGCGGCATCTGA 193
|||
QY 135 TTTGTCCCTGGGGCGGCGAGCGCGGAGATGAGGCGGTGATTTAGCAAGTTAA 194
|||
Db 194 TTTGTCCCTGGGGCGGCGAGCGCGGAGATGAGGCGGTGATTTAGCAAGTTAA 253
|||
QY 195 AAGTAACGAACCATGGCTCAGTTCCAAACCTTTTGTGGCAGCCTGGAATATCTGGGC 254
|||
Db 254 AAGTAACGAACCATGGCTCAGTTCCAAACCTTTTGTGGCAGCCTGGAATATCTGGGC 313
|||
QY 255 CATTAAGTAGAGGAAAGAGCGAAGCATGATCAGCAGTTCCATAGTTTAAAGCCATATC 314
|||
Db 314 CATTAAGTAGAGGAAAGAGCGAAGCATGATCAGCAGTTCCATAGTTTAAAGCCATATC 373
|||
QY 315 TGGATTCATTAAGTGTGATCAAGCTAGAACTTTTTCATCTGGGTTACCTCAACG 374
|||
Db 374 TGGATTCATTAAGTGTGATCAAGCTAGAACTTTTTCATCTGGGTTACCTCAACG 433
|||
QY 375 TGTTTAGCACAGATATGGGCACTAGTACATGATATGATGAGAGATGATCAAGT 434
|||
Db 434 TGTTTAGCACAGATATGGGCACTAGTACATGATATGATGAGAGATGATCAAGT 493
|||
QY 435 GGAAGTTTCCATAGCTATGAAGCTTATCAAACTGAAGCTACAGAGATTCAGTACCCTC 494
|||
Db 494 GGAAGTTTCCATAGCTATGAAGCTTATCAAACTGAAGCTACAGAGATTCAGTACCCTC 553
|||
QY 495 TGGACT 500
|||
Db 554 TGGACT 559
```

Search completed: July 1, 2004, 12:16:42
Job time : 2138 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2004, 10:19:06 ; Search time 2656 Seconds

(without alignments)
5621.645 Million cell updates/sec

Title: US-09-720-934-1_COPY_1_500

Perfect score: 500
Sequence: 1 caaagaatccgggtacg.....taccagctaccctctgact 500

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estlum:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_fyt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_pro:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	486	97.2	896	13	BQ941336 AGENCOURT
2	486	97.2	1089	12	BQ050397 AGENCOURT
3	484.4	96.9	979	13	BQ941411 AGENCOURT
4	474	94.8	750	12	BG118422 AGENCOURT

5	474	94.8	2126	11	BC020269 Homo sapi
6	460	92.0	643	13	BX470886 DKFZp686E
7	459.8	92.0	811	12	BG829540 G02763842
8	454	90.8	544	14	AL711737 DKFZp686K
9	437.4	87.5	605	14	CB296595 12822043
10	424	84.8	2079	11	BC013578 Homo sapi
11	413	82.6	734	14	CB269754 1008661 H
12	397.6	79.5	405	9	AL750406 cno2G02.x
13	387.6	77.5	547	10	BF194430 246586 MA
14	367.4	73.5	698	14	CB519152 UI-M-GH0-
15	367.4	73.5	5385	11	BC062938 Mus muscu
16	366.2	73.2	723	10	BE786696 601475048
17	363.8	72.8	624	13	BY713771 BY713771
18	329.6	65.9	662	14	CE536420 UI-M-GT0-
19	323.2	64.6	600	14	CA528645 8082-31 M
20	298.2	59.6	600	14	CA529211 8101-57 M
21	297	59.4	411	9	AL549192 UI-R-C3-t
22	293.2	58.6	475	10	BE199657 ug70a11.x
23	293.2	58.6	488	9	AL481705 v330f01.x
24	292.2	58.4	445	9	AL852079 UI-M-BH0-
25	292.2	58.4	445	9	AL852070 UI-M-BH0-
26	292.2	58.4	501	10	BE952536 UI-M-CB0-
27	291.8	58.4	394	12	BM484041 UI-M-BH0-
28	289	57.8	359	9	AL853008 UI-M-BH0-
29	283.2	56.6	452	10	BB840012 BB840012
30	276.8	55.4	523	13	B0381021 603860857
31	271.4	54.3	533	29	CG579249 OSM217389
32	270	54.0	271	13	EX470913 DKFZp686G
33	267.4	53.5	728	13	BY731854 BY731854
34	266.4	53.3	766	14	CF743758 UI-M-GI0-
35	254	50.8	477	14	CB269094 1008001 H
36	249.8	50.0	907	13	B0128229 603114019
37	246.8	49.4	873	14	CF728634 UI-M-HD0-
38	245.2	49.0	298	10	BE103413 UI-R-BX0-
39	244.6	48.9	299	10	AM890668 QV1-NT004
40	232.8	46.6	643	13	BU406611 604136611
41	218.2	43.6	589	9	AL792490 AL792490
42	216.6	43.3	534	12	BG160235 de42d08.Y
43	215.6	43.1	550	12	B0623100 B0623100
44	215	43.0	504	9	AL964682 AL964682
45	196.6	39.3	266	10	BE954163 UI-M-CD1-

ALIGNMENTS

RESULT 1
LOCUS BQ941336
DEFINITION AGENCOURT_8741326 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6420600
5', mRNA sequence.
ACCESSION BQ941336
VERSION BQ941336.1 GI:22356814
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Metazoa; Chordata; Vertebrata; Euteleostomi;
Eukaryota; Euteleostomi; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: DCTD/DP/GenDat
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLN at:
http://image.llnl.gov
Plate: LCM2594 row: k column: 01
High quality sequence stop: 763.

FEATURES
source
Location/Qualifiers
1..896
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5420600"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_18"
/note="Organ: Lung; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

ORIGIN
Query Match 97.2%; Score 486; DB 13; Length 896;
Best Local Similarity 100.0%; Pred. No. 7.3e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GTACGGCGGCTCGCAGAGAAATCCCGAGCGGCTCCGGAGCGACAGAGAGCGGGCG 74
DB 17 GTACGGCGGCTCGCAGAGAAATCCCGAGCGGCTCCGGAGCGAGAGAGAGCGGGCG 76
QY 75 GGATGTGTGTGGGGGCTGGGCTCTGCTGCTCCAGCGGCGGTGAGCGGCACTGA 134
DB 77 GGATGTGTGTGGGGGCTGGGCTCTGCTGCTCCAGCGGCGGTGAGCGGCACTGA 136
QY 135 TTGTGCTGTGGGGGCGCAGCGCGGCGGAGATGAGCGGTGAGCGGCACTGA 194
DB 137 TTGTGCTGTGGGGGCGCAGCGCGGCGGAGATGAGCGGTGAGCGGCACTGA 196
QY 195 AAGTAAAGAGCAATGGCTCAATTTTCCAAACACCTTTTGGGAGCGCTGATATCTGGG 254
DB 197 AAGTAAAGAGCAATGGCTCAATTTTCCAAACACCTTTTGGGAGCGCTGATATCTGGG 256
QY 255 CATTAAGTGAAGAGAGAGAGAGAGATGATGAGCGGTGAGCGGCACTGA 314
DB 257 CATTAAGTGAAGAGAGAGAGAGATGATGAGCGGTGAGCGGCACTGA 316
QY 315 TGGATTCATTACTGTGATCAAGCTAGAACTTTTTCATCTGGGTTACTCAACC 374
DB 317 TGGATTCATTACTGTGATCAAGCTAGAACTTTTTCATCTGGGTTACTCAACC 376
QY 375 TGTTTAGACAGATATGGGCACTAGCTGACATGAATATGATGAGAGATGATCAAGT 434
DB 377 TGTTTAGACAGATATGGGCACTAGCTGACATGAATATGATGAGAGATGATCAAGT 436
QY 435 GGAGTTTCCATAGCATGAATGAATCTTCAAACTGAGAGCTCAAGAGATTCAGTCAACC 494
DB 437 GGAGTTTCCATAGCATGAATGAATCTTCAAACTGAGAGCTCAAGAGATTCAGTCAACC 496
QY 495 TGCACACT 500
DB 497 TGCACACT 502

RESULT 2
BQ050397 1089 bp mRNA linear EST 29-MAR-2002
LOCUS AGENCOURT_7050803 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5784342
DEFINITION 5' mRNA Sequence.
ACCESSION BQ050397
VERSION BQ050397.1 GI:19809737
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1089)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Plate: LLM12869 row: 1 column: 07
High quality sequence stop: 612.

FEATURES
source
Location/Qualifiers
1..1089
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5784342"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."

ORIGIN
Query Match 97.2%; Score 486; DB 12; Length 1089;
Best Local Similarity 100.0%; Pred. No. 7.7e-116;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GTACGGCGGCTCGCAGAGAAATCCCGAGCGGCTCCGGAGCGACAGAGAGCGGGCG 74
DB 57 GTACGGCGGCTCGCAGAGAAATCCCGAGCGGCTCCGGAGCGACAGAGAGCGGGCG 116
QY 75 GGATGTGTGTGGGGGCTGGGCTCTGCTGCTCCAGCGGCGGTGAGCGGCACTGA 134
DB 117 GGATGTGTGTGGGGGCTGGGCTCTGCTGCTCCAGCGGCGGTGAGCGGCACTGA 176
QY 135 TTGTGCTGTGGGGGCGCAGCGCGGCGGAGATGAGCGGTGAGCGGCACTGA 194
DB 177 TTGTGCTGTGGGGGCGCAGCGCGGCGGAGATGAGCGGTGAGCGGCACTGA 226
QY 195 AAGTAAAGAGCAATGGCTCAATTTTCCAAACACCTTTTGGGAGCGCTGATATCTGGG 254
DB 237 AAGTAAAGAGCAATGGCTCAATTTTCCAAACACCTTTTGGGAGCGCTGATATCTGGG 296
QY 255 CATTAAGTGAAGAGAGAGAGAGATGATGAGCGGTGAGCGGCACTGA 314
DB 297 CATTAAGTGAAGAGAGAGAGATGATGAGCGGTGAGCGGCACTGA 356
QY 315 TGGATTCATTACTGTGATCAAGCTAGAACTTTTTCATCTGGGTTACTCAACC 374
DB 357 TGGATTCATTACTGTGATCAAGCTAGAACTTTTTCATCTGGGTTACTCAACC 416
QY 375 TGTTTAGACAGATATGGGCACTAGCTGACATGAATATGATGAGAGATGATCAAGT 434
DB 417 TGTTTAGACAGATATGGGCACTAGCTGACATGAATATGATGAGAGATGATCAAGT 476
QY 435 GGAGTTTCCATAGCATGAATGAATCTTCAAACTGAGAGCTCAAGAGATTCAGTCAACC 494
DB 477 GGAGTTTCCATAGCATGAATGAATCTTCAAACTGAGAGCTCAAGAGATTCAGTCAACC 536
QY 495 TGCACACT 500
DB 537 TGCACACT 542

RESULT 3
BQ941411 979 bp mRNA linear EST 21-AUG-2002
LOCUS BQ941411
DEFINITION AGENCOURT_8744157 lupskl_sciatic nerve Homo sapiens cDNA clone

IMAGE:6205905 5', mRNA sequence.
ACCESSION B0941411
VERSION B0941411.1 GI:22356889
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 979)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLM13629 row: 1 column: 10
High quality sequence stop: 599.
Location/Qualifiers
1..979
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6205905"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="Lupski sciatic nerve"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors:
5'-TCGACCCACGCTCCG-3' and
5'-GACTAGTCTAGATCGAGCGGCGCCCTT(15)-3'. Size selected > 1 kb for average insert length 1.87 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."
ORIGIN
Query Match 96.9%; Score 484.4; DB 13; Length 979;
Best Local Similarity 99.8%; Pred. No. 2e-115;
Matches 485; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
15 GTACGGCGGCTCGCGAGAGAAATCCGAGCGGCTTCGGGACGACAGAGGCGGCG 74
Db GTACGGCGGCTCGCGAGAGAAATCCGAGCGGCTTCGGGACGACAGAGGCGGCG 152
QY GGGATGATGTCGGGCGCTCGGCTCCTGCTCCGAGCGGCGGTGAGCGGCACTGA 134
Db GGGATGATGTCGGGCGCTCGGCTCCTGCTCCGAGCGGCGGTGAGCGGCACTGA 212
QY TTTGTCCCTGGGCGCGACCGCGCCCGGAGATGAGCGTCAATTAGCAAGTAA 194
Db TTTGTCCCTGGGCGCGACCGCGCCCGGAGATGAGCGTCAATTAGCAAGTAA 272
QY AAGTAAAGAACATGAGCTCACTTTTCCAAACCTTTTGGTGGAGCTGATATCTGGGC 254
Db AAGTAAAGAACATGAGCTCACTTTTCCAAACCTTTTGGTGGAGCTGATATCTGGGC 332
QY CATACTGAGAGGAAAGAGGAGATGATGAGATGATGATGATGATGATGATGATGATG 314
Db CATACTGAGAGGAAAGAGGAGATGATGAGATGATGATGATGATGATGATGATGATG 392
QY TGGATTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 374
Db TGGATTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 452

QY 375 TGGTTAGCAGATATGGGACAGTACATGATGATGATGATGATGATGATGATGATGAT 434
Db 453 TGGTTAGCAGATATGGGACAGTACATGATGATGATGATGATGATGATGATGATGAT 512
QY 435 GGAGTTTCCATAGCTATGAACTTATCAACTGAGCTAGACAGATATCACTACCTC 494
Db 513 GGAGTTTCCATAGCTATGAACTTATCAACTGAGCTAGACAGATATCACTACCTC 572
QY 495 TGCACCT 500
Db 573 TGCACCT 578
RESULT 4
B0118422 750 bp mRNA linear EST 30-JAN-2001
LOCUS 602348040F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:444129 5',
DEFINITION mRNA sequence.
ACCESSION B0118422 GI:12611928
VERSION B0118422.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 750)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLM10216 row: p column: 10
High quality sequence stop: 642.
Location/Qualifiers
1..750
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:444129"
/tissue_type="adenocarcinoma, cell line"
/clone_lib="NIH_MGC_90"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
ORIGIN
Query Match 94.8%; Score 474; DB 12; Length 750;
Best Local Similarity 99.8%; Pred. No. 9.4e-113;
Matches 485; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
15 GTACGGCGGCTCGCGAGAGAAATCCGAGCGGCTTCGGGACGACAGAGGCGGCGG 74
Db 101 GTACGGCGGCTCGCGAGAGAAATCCGAGCGGCTTCGGGACGACAGAGGCGGCGG 160
QY GGGATGATGTCGGGCGCGACCGCGCCCGGAGATGAGCGTCAATTAGCAAGTAA 134
Db 161 GGGATGATGTCGGGCGCGACCGCGCCCGGAGATGAGCGTCAATTAGCAAGTAA 220
QY TTTGTCCCTGGGCGCGACCGCGCCCGGAGATGAGCGTCAATTAGCAAGTAA 194
Db 221 TTTGTCCCTGGGCGCGACCGCGCCCGGAGATGAGCGTCAATTAGCAAGTAA 280

QY 195 AAGTAAACAGACCATGGCTCACTTTCCACACCTTTGGTGAGCGCTGGATATCTGGG 254
| | | | |
Db 281 AATTAACAGACCATGGCTCACTTTCCACACCTTTGGTGAGCGCTGGATATCTGGG 340
| | | | |
QY 255 CATTAAGTGAAG 314
| | | | |
Db 341 CATTAAGTGAAG 400
| | | | |
QY 315 TGGATCATTAAGTGAAG 374
| | | | |
Db 401 TGGATCATTAAGTGAAG 460
| | | | |
QY 375 TGGTGAAG 434
| | | | |
Db 461 TGGTGAAG 520
| | | | |
QY 435 GGAAGTTTCCATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 494
| | | | |
Db 521 GGAAGTTTCCATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 579
| | | | |
QY 495 TGGACT 500
| | | | |
Db 580 TGGACT 585

RESULT 5
BC020269 2126 bp mRNA linear HTC 19-DEC-2001
LOCUS Homo sapiens, clone IMAGE:4899011, mRNA.
DEFINITION BC020269
ACCESSION BC020269.1 GI:17939664
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1 (bases 1 to 2126)
JOURNAL Strausberg, R.
Direct Submission
Submitted (19-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov

Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdel, Yaron Butterfield,
Susanna Chan, Readman Chin, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywicki, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Matheson, Candace McLeary, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schin, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 40 Row: n Column: 1
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4504796
This clone has the following problem: no cloning site /
microdeletion.

FEATURES
Source
1. 2126
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 94.8%; Score 474; DB 11; Length 2126;
Best Local Similarity 100.0%; Pred. No. 1.2e-112;
Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/clone="IMAGE:4899011"
/tissue_type="Pancreas, epithelioid carcinoma"
/clone_id="NIH_MGC_42"
/lab_host="DH10B-R"
/note="Vector: pORF7"

QY 27 GCGAGAGAGATCCCGAGGAGGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 86
| | | | |
Db 1 GCGAGAGAGATCCCGAGGAGGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
| | | | |
QY 87 GGGGCTGGGCTCTCTGCTCTCTCCAGAGGCGCGCTGAGCGGCACTGATTTGCTGGG 146
| | | | |
Db 61 GGGGCTGGGCTCTCTGCTCTCTCCAGAGGCGCGCTGAGCGGCACTGATTTGCTGGG 120
| | | | |
QY 147 GCGGAGCGGCGAGCGCGCGCGAGATGAGGCGCTGATGAGAGATGAGAGAGAGAGAG 206
| | | | |
Db 121 GCGGAGCGGCGAGCGCGCGCGAGATGAGGCGCTGATGAGAGAGATGAGAGAGAGAG 180
| | | | |
QY 207 CATGCTCAGTTTCCACACCTTTTGGTGGCAGCTGATATCTGGGCGCATTAAGTGA 266
| | | | |
Db 181 CATGCTCAGTTTCCACACCTTTTGGTGGCAGCTGATATCTGGGCGCATTAAGTGA 240
| | | | |
QY 267 GGAAG 326
| | | | |
Db 241 GGAAG 300
| | | | |
QY 327 TGGTGAAG 386
| | | | |
Db 301 TGGTGAAG 360
| | | | |
QY 387 GATATGGGAG 446
| | | | |
Db 361 GATATGGGAG 420
| | | | |
QY 447 AGCTATGAAACTTATCAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 500
| | | | |
Db 421 AGCTATGAAACTTATCAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 474
| | | | |

RESULT 6
BX470886

LOCUS BX470886 643 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp686E17123_r1 686 (synonym: h1cc3) Homo sapiens CDNA clone
ACCESSION DKFZp686E17123.5, mRNA sequence.
BX470886

KEYWORDS BX470886.1 GI:31665209

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 643)

AUTHORS Bahr, A., Lauber, J., Mewes, H.W., Well, B., Amlid, C., Osanger, A.,
Robo, G., Han, W., and Wiemann, S.
EST (Bahr, A., Lauber, J., Mewes, H.W., Well, B., et al.)
Unpublished (2003)
Contact: MIPS

TITLE MIPS

Ingolstraedter Landstr.1, D-85764 Neuberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de,
sequenced by Qiagen (Hilden/Germany) within the CDNA sequencing
consortium of the German Genome Project.

No sl sequence available.
This clone (DKFZp686E17123) is available at the RZPD in Berlin.
Please contact the RZPD: Reesourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

Source

Location/Qualifiers

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKR2686R17123"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="666 (synonym: hicc3)"
/name="Vector: pTribLEX2; Site_1: SfiI; Site_2: SfiIB
      DNACollection"

```

ORIGIN

Query Match	92.0%;	Score 460;	DB 13;	Length 643;
Best Local Similarity	99.8%;	Pred. No. 4.1e-109;		
Matches 471; Conservative	0;	Mismatches 0;	Indels 1;	Gaps 1;

QY	29	GAGGAAGAAATCCCGACGGGCTCGGGACGGACAGAGAGGCGGGCGGGAAATGATGTGCGG	88
Db	2	GAGGAAGAAATCCCGACGGGCTCGGGACGGACAGAGAGGCGGGCGGGAAATGATGTGCGG	61
QY	89	GGCTCGGCTCCTGTGCTCCTCCACGGCGCGGTGAGCGCATGTATTTGTCCTCGGAGC	148
Db	62	GGCTCGGCTCCTGTGCTCCTCCACGGCGCGGTGAGCGCATGTATTTGTCCTCGGAGC	121
QY	149	GGCAGCGCGGACCGGCGCGAGATGAGGCGTGCATTACAGAGGTAAAGTAAACAACA	208
Db	122	GGCAGCGCGGACCGGCGCGAGATGAGGCGTGCATTACAGAGGTAAAGTAAACAACA	181
QY	209	TGCGTCAGTTTCGAACACCTTTTGGTGGCAGCGCTGGATATCTGGGCCATAACTGTAGAG	268
Db	182	TGCGTCAGTTTCGAACACCTTTTGGTGGCAGCGCTGGATATCTGGGCCATAACTGTAGAG	241
QY	269	AAAGAGCGAAGCATGATGACGAGTTCCATGTTTAAAGCCAAATATCTGGATTCATTACTG	328
Db	242	AAAGAGCGAAGCATGATGACGAGTTCCATGTTTAAAGCCAAATATCTGGATTCATTACTG	301
QY	329	GTGATCAAGCTAGAAACCTTTTTTTTCAATCTGGGTTCCTCAACTGTGTTTAGACACAGA	388
Db	302	GTGATCAAGCTAGAAACCTTTTTTTTCAATCTGGGTTCCTCAACTGTGTTTAGACACAGA	360
QY	389	TATGGGCACTAGCTGACATGAAATAATGATGAGAAATGGATCAAGTGAAGTTCATAG	448
Db	361	TATGGGCACTAGCTGACATGAAATAATGATGAGAAATGGATCAAGTGAAGTTCATAG	420
QY	449	CTATGAACCTATCAAACTGAAGCTACAAAGATATACGATACCCTCTGCACCT	500
Db	421	CTATGAACCTATCAAACTGAAGCTACAAAGATATACGATACCCTCTGCACCT	472

RESULT	7
BG829540	
LOCUS	
DEFINITION	BG829540 811 bp mRNA EST 22-MAY-2006 602763842P1 NIH_MGC_42 Homo sapiens CDNA clone IMAGE:4895011 5', mRNA sequence.
ACCESSION	BG829540
VERSION	BG829540.1 GI:14177114
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 811) NIH-MGC http://mgc.nci.nih.gov/ .
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE	Unpublished (1999)
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cga@bs-remail.nih.gov
COMMENT	Tissue Procurement: ARCC cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution information can be found at:

FEATURES

source

found through the I.M.A.G.E. Consortium/LINL at
http://image.llnl.gov
Plate: L1CM1790 row: c column: 12
High quality sequence stop: 651.

```

1.811
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4699011"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH MGC 42"
/notes="Organ: pancreas; Vector: pOT87; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAC(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Straatene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library. |"

```

ORIGIN

Query Match	92.0%;	Score 459.8;	DB 12;	Length 811;
Best Local Similarity	99.4%;	Pred. No. 4.9e-109;		
Matches 472; Conservative	0;	Mismatches 2;	Indels 1;	Gaps 1;

Oy	27	GCAGAGAAAGATCCCGAGCGGGCTCCCGGACGAGACAGAGAGCGCGGCGAGATGGTGGC	86
Db	2	GCAGAGAAAGATCCCGAGCGGGCTCCCGGACGAGACAGAGAGCGCGGCGAGATGGTGGC	61
Oy	87	GGGGCTGCGGCTCTGTGGCTCCCGACGCGCGCGTGAAGCGGCACTGATTTGTCCCTGGG	146
Db	62	GGGGCTGCGGCTCTGTGGCTCCCGACGCGCGCGTGAAGCGGCACTGATTTGTCCCTGGG	121
Oy	147	GGCGACACGGGACCCCGCCGAGATAGAGGCTGATTAGCAAGGTAAAGTAAACAGAAC	206
Db	122	GGCGACACGGGACCCCGCCGAGATAGAGGCTGATTAGCAAGGTAAAGTAAACAGAAC	181
Oy	207	CATGGCTCAGTTTCCAAACACCTTTTGGTG-6CAGCCTGGATATCTGGGCCATTAACCTGTAG	265
Db	182	CATGGCTCAGTTTCCAAACACCTTTTGGTGCGACGCTGGATATCTGGGCCATTAACCTGTAG	241
Oy	266	AGGAAAGAGCGAAGCATGATAGACAGTTCCTAGTTTAAAGCCCAATATCTGGATTCATTA	325
Db	242	AGGAAAGAGCGAAGCATGATAGACAGTTCCTAGTTTAAAGCCCAATATCTGGATTCATTA	301
Oy	326	CTGGTGATCAAGCTAGAGAACTTTTTCATCTGGGTTACCTCAACCTGTTTTAGACAC	385
Db	302	CTGGTGATCAAGCTAGAGAACTTTTTCATCTGGGTTACCTCAACCTGTTTTAGACAC	361
Oy	386	AGATATGGGACTAGCTGACATGATATATGATGAGAGATGATCAAGTGAATTTTCCA	445
Db	362	AGATATGGGACTAGCTGACATGATATATGATGAGAGATGATCAAGTGAATTTTCCA	421
Oy	446	TAGCTATCAAACTATCAAACTGAAGGTACAGAGATTAACACTACCTCCCTGGACT	500
Db	422	TAGCTATCAAACTATCAAACTGAAGGTACAGAGATTAACACTACCTCCCTGGACT	476

RESULT	8			
LOCUS	AL711737			
DEFINITION	AL711737	544 bp	mRNA	linear EST 04-SBP-2003
ACCESSION	DKEZP686K1884	r1.686	(synonym: hlc3)	Homo sapiens cDNA clone
VERSION	DKEZP686K1884.5.			
KEYWORDS	AL711737			
SOURCE	AL711737.1	GI:19695092		
ORGANISM	Homo sapiens (human)			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			

REFERENCE 1 (bases 1 to 544)
 AUTHORS Duesterhoeft, A., Lauber, J., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
 TITLE EST (Duesterhoeft, et al.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: MIPS
 MIPS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 No 5' sequence available.
 This clone (DKFZ6686184) is available at the RZPD in Berlin.
 Please contact the RZPD; Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

1..544
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZ6686184"
 /dev_stage="adult"
 /lab_host="DH1B"
 /note="Vector: pTRIPLEX2; Site_1: SfiIA; Site_2: SfiIB;
 cDNA-collection"

ORIGIN

Query Match 90.8%; Score 454; DB 9; Length 544;
 Best Local Similarity 99.8%; Pfed. No. 1.4e-107;
 Matches 465; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

35 GATCCGAGCGGCGCTCCGGACGCGACAGAGAGCGCGGAGTGTGTGCGGGGCTGC 94
 |||||
 2 GATCCGAGCGGCGCTCCGGACGCGACAGAGAGCGCGGAGTGTGTGCGGGGCTGC 61
 |||||
 95 GGCTCCTGCGTCCCTCCAGCGGCGGTGAGCGGCGGCGGCGGCGGCGGCGGCGG 154
 |||||
 62 GGCTCCTGCGTCCCTCCAGCGGCGGTGAGCGGCGGCGGCGGCGGCGGCGGCGG 120
 |||||
 155 GCGGACCGCGCGGAGATGAGCGGTGATTAGCAAGTAAAGTAAAGTAAAGTAAAG 214
 |||||
 121 GCGGACCGCGCGGAGATGAGCGGTGATTAGCAAGTAAAGTAAAGTAAAGTAAAG 180
 |||||
 215 AGTTTCCACACCTTTGGTGGGCGGCGGTGATCTGGGCGCATTAAGTAAAGTAAAG 274
 |||||
 181 AGTTTCCACACCTTTGGTGGGCGGCGGTGATCTGGGCGCATTAAGTAAAGTAAAG 240
 |||||
 275 CGAAGCATGATGACGAGTTCATAGTTTAAAGCCAAATATCTGGATTCATTACTGTGATC 334
 |||||
 241 CGAAGCATGATGACGAGTTCATAGTTTAAAGCCAAATATCTGGATTCATTACTGTGATC 300
 |||||
 335 AAGCTTGAACCTTTTTCATCTGGGTTACCTCAACCTGTTTAAAGCAGATATGGG 394
 |||||
 301 AAGCTTGAACCTTTTTCATCTGGGTTACCTCAACCTGTTTAAAGCAGATATGGG 360
 |||||
 395 CACTAGCTGACATGATATGATGAAAGATGATCAAGTGGAGTTTCATAGTATGATGA 454
 |||||
 361 CACTAGCTGACATGATATGATGAAAGATGATCAAGTGGAGTTTCATAGTATGATGA 420
 |||||
 455 AACTATCAACTGAAGCTACAAAGATATCAAGTACCTCTGCACT 500
 |||||
 421 AACTATCAACTGAAGCTACAAAGATATCAAGTACCTCTGCACT 466
 |||||

RESULT 9
 CB296595 605 bp mRNA linear EST 28-FEB-2003
 LOCUS 12B22043 rev 1 A05 r 037.abi Chimpanzee brain library Kooos Pan
 DEFINITION tlog10bytes cDNA clone 12B22043_rev_1_A05_r_037.abi 5', mRNA
 sequence.

ACCESSION CB296595
 VERSION CB296595.1 GI:28622025
 KEYWORDS EST.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes

REFERENCE 1 (bases 1 to 605)
 AUTHORS Hellmann, I., Zollner, S., Enard, W., Ebersberger, I., Nickel, B. and Paabo, S.
 TITLE Selection on human genes as revealed by comparisons to chimpanzee
 JOURNAL Genome Res. (2003) In press
 COMMENT Contact: Paabo S
 Evolutionary Genetics
 Max-Planck-Institute for evolutionary Anthropology
 Deutscher Platz 6, 04103 Leipzig, Germany
 Tel: +49-(0)-341-3550 500
 Fax: +49-(0)-341-3550 555
 Email: paabo@eva.mpg.de
 Seq primer: M13 reverse.

FEATURES

source

1..605
 /organism="Pan troglodytes"
 /mol_type="mRNA"
 /db_xref="taxon:9598"
 /clone="12B22043_rev_1_A05_r_037.abi"
 /sex="male"
 /tissue_type="brain, presumably cortex"
 /dev_stage="adult"
 /lab_host="Episcurian Coli (TM) Xu-10-Gold"
 /clone_lib="Chimpanzee brain library Kooos"
 /note="Vector: pUC19; Site_1: SfiI-A; Site_2: SfiI-B; The
 library was prepared using the SMART cDNA library
 construction kit (Clontech), doing only primer extension,
 but not PCR amplification of the cDNA. The only deviation
 from the published protocol was that we cloned the cDNA
 into a plasmid vector."

ORIGIN

Query Match 87.5%; Score 437.4; DB 14; Length 605;
 Best Local Similarity 99.6%; Pfed. No. 3.2e-103;
 Matches 449; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

50 TCCGGGACGAGACAGAGAGCGGCGGAGTGGTGGCGGCGGCGGCGGCGGCGGCTC 109
 |||||
 1 TCCGGGACGAGACAGAGAGCGGCGGAGTGGTGGCGGCGGCGGCGGCGGCGGCTC 60
 |||||
 110 CCCAGCGGCGGCGTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 169
 |||||
 61 CCCAGCGGCGGCGTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 |||||
 170 GATGAGCGCGTGCATATGCAAGGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 229
 |||||
 121 GATGAGCGCGTGCATATGCAAGGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 180
 |||||
 230 TTGGTGGCAGCGCTGATATCTGGGCCCATTAAGTAAAGTAAAGTAAAGTAAAGTAA 289
 |||||
 181 TTGGTGGCAGCGCTGATATCTGGGCCCATTAAGTAAAGTAAAGTAAAGTAAAGTAA 240
 |||||
 290 AGTTCCATGTTTAAAGCCAAATATCTGATTTCACTGATTTCAAGTAAAGTAAAGTAA 349
 |||||
 241 AGTTCCATGTTTAAAGCCAAATATCTGATTTCACTGATTTCAAGTAAAGTAAAGTAA 299
 |||||
 350 TTTTTCATCTGGTTACTCAACCGTTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 409
 |||||
 300 TTTTTCATCTGGTTACTCAACCGTTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 359
 |||||
 410 ATATATGATGAAGATGATGATCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 469
 |||||
 360 ATATATGATGAAGATGATGATCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 419
 |||||
 470 AGCTACAGAGATATACGATTAACCTCTGCACT 500
 |||||

DB 420 AGCTAAGAGATATCAGCTACCCCTGCACT 450

RESULT 10
BC013578
LOCUS
DEFINITION Homo sapiens, Similar to Intersectin 1 (SH3 domain protein), clone IMAGE:3878242, mRNA.
ACCESSION BC013578
VERSION BC013578.1 GI:15488896
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2079)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (04-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapds@mail.nih.gov
Tissue Procurement: DCTD/DRP/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNI)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcdpaxil@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLNI at: <http://image.llnl.gov>
Series: IRAX Plate: 14 Row: C Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4504796
This clone has the following problem: retained intron.
FEATURES
Location/Qualifiers
1..2079
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3878242"
/tissue_type="lung, large cell carcinoma"
/clone_lib="NTH MGC_68"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN
Query Match 84.8%; Score 424; DB 11; Length 2079;
Best Local Similarity 100.0%; Pred. No. 1.4e-99;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 77 GATGGTGTGGGGGCTGGGGCTCTGCGTCTCCAGCGGGCGCTGAGCGGCACTGATT 136
1 GATGTTGTGGGGGCTGGGGCTCTGCGTCTCCAGCGGGCGCTGAGCGGCACTGATT 60
QY 137 TGTCCCTGGGGCGGAGCGCGGACCCGCCGAGATGAGGCGTGAATTAGCAAGGTAA 196
61 TGTCCCTGGGGCGGAGCGCGGACCCGCCGAGATGAGGCGTGAATTAGCAAGGTAA 120
DB 197 GTAAGAGACCATGGCTCACTTTCACACACCTTTTGTGGAGCGCTGATATCTGGGCA 256
121 GTAAGAGACCATGGCTCACTTTCACACACCTTTTGTGGAGCGCTGATATCTGGGCA 180
QY 257 TAACTGTAGAGAGAGAGAGAGATGATCAGACAGTTCATAGTTAAAGCCATATCTG 316
181 TAACTGTAGAGAGAGAGAGAGATGATCAGACAGTTCATAGTTAAAGCCATATCTG 240

QY 317 GATTCATTACTGTGTATCAAGCTTGAACCTTTTTCANTCTGGGTTACTCAACTG 376
DB 241 GATTCATTACTGTGTATCAAGCTTGAACCTTTTTCANTCTGGGTTACTCAACTG 300
QY 377 TTTTAGACACAGATATNGGGCACTAGCTGACATGAATTAAGATGAAGATGATCAAGTG 436
DB 301 TTTTAGACACAGATATNGGGCACTAGCTGACATGAATTAAGATGAAGATGATCAAGTG 360
QY 437 AGTTTCCATAGCTATGAAACTTATCAAACTGAAGCTCAAGAGATATCAAGTCCCTG 496
DB 361 AGTTTCCATAGCTATGAAACTTATCAAACTGAAGCTCAAGAGATATCAAGTCCCTG 420
QY 497 CACT 500
DB 421 CACT 424

RESULT 11
CB269754
LOCUS 734 bp mRNA linear EST 20-FEB-2003
DEFINITION 1008661 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens
CDNA 5', mRNA sequence.
ACCESSION CB269754
VERSION CB269754.1 GI:28444339
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 734)
AUTHORS Yang, R.-Z., Shuldiner, A. and Gong, D.-W.
TITLE EST analysis of human adipose gene expression
JOURNAL Unpublished (2002)
COMMENT Contact: Gong Da-Wei
Division of Endocrinology, Diabetes and Nutrition
University of Maryland
660 Redwood St, BH497, Baltimore, MD 21201, USA
Tel: 410 706 1672
Fax: 410 706 1622
Email: dgon@medicine.umaryland.edu
PCR Primers
FORWARD: CTCGGAGAGCGCGCATTTGTGTTGCT
BACKWARD: AATACGACTCATATGAGGCGCAATTGG
Seq primer: GTTGGTACCGGAATTC.
FEATURES
Location/Qualifiers
1..734
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="Male and Female"
/tissue_type="Adipose"
/clone_lib="Human Fat Cell 5'-Stretch Plus cDNA library"
/note="Vector: lambda triplex"

ORIGIN
Query Match 82.6%; Score 413; DB 14; Length 734;
Best Local Similarity 100.0%; Pred. No. 7.8e-97;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 88 GGGCTGGGGCTCTGCGTCTCCAGCGGGCGCTGAGCGGCACTGATTGCTCGGG 147
1 GGGCTGGGGCTCTGCGTCTCCAGCGGGCGCTGAGCGGCACTGATTGCTCGGG 60
QY 148 CGGACGCGGAGACCGCGCGAGATGAGCGTGAATTAGCAAGTAAAGTAAAGAAC 207
61 CGGACGCGGAGACCGCGCGAGATGAGCGTGAATTAGCAAGTAAAGTAAAGAAC 120
DB 208 ATGCTCAGTTTCAACACCTTTTGTGGAGCGCTGATATCTGGGCCATTACTGTAG 267
121 ATGCTCAGTTTCAACACCTTTTGTGGAGCGCTGATATCTGGGCCATTACTGTAG 180
QY 268 GAAAGAGGAGAGATGATCAGACAGTTCATAGTTAAAGCCATATCTGATCTACT 327

Db 181 GAAAGAGCGAGCATGATGAGCAGTTCATAGTTTAAAGCCATATCTGATTCATTACT 240
 QY 328 GGGATCAAGCTTGAAGAACTTTTTCATCTGGGTTTACTCAACTGTTTATAGCAG 387
 Db 241 GGGATCAAGCTTGAAGAACTTTTTCATCTGGGTTTACTCAACTGTTTATAGCAG 300
 QY 388 ATATGGGCACTAGCTAGCATGATTAATGATGAGAAGATGATCAAGTGAAGTTTCCATA 447
 Db 301 ATATGGGCACTAGCTAGCATGATTAATGATGAGAAGATGATCAAGTGAAGTTTCCATA 360
 QY 448 GCTATGAACCTTATCAAACTGAAAGCTTCAAGATATCAAGTCACTGCTGCACT 500
 Db 361 GCTATGAACCTTATCAAACTGAAAGCTTCAAGATATCAAGTCACTGCTGCACT 413

RESULT 12
 A1750406 405 bp mRNA linear EST 20-JUN-2002
 LOCUS cno2g02.x2 Normal Human Trabecular Bone Cells Homo sapiens cDNA
 DEFINITION clone NHTBC_cno2g02 random, mRNA sequence.
 ACCESSION A1750406
 VERSION A1750406.1 GI:5128670
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 405)
 Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G., Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M., Robey,P.G., Hotchkiss,R.N. and Francomano,C.A.
 SCAP: The Skeletal Genome Anatomy Project
 Unpublished (1997)

TITLE
 JOURNAL Contact: Libin Jia
 COMMENT Medical Genetics Branch
 National Human Genome Research Institute
 10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
 Tel: 301-402-4877
 Fax: 301-496-7157
 Email: libin@helix.nih.gov

FEATURES
 source
 1.405
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="NHTBC_cno2g02"
 /sex="female"
 /tissue_type="bone"
 /cell_type="Trabecular Bone Cells"
 /lab_host="SURE"
 /clone_lib="Normal Human Trabecular Bone Cells"
 /note="Organ: Hip; Vector: pBluescript; Site 1: EcoRI;
 library constructed by Dr. Marian Young and Dr. Pamela Gehron Robey (NIDCR)"

ORIGIN
 Query Match 79.5%; Score 397.6; DB 9; Length 405;
 Best Local Similarity 99.0%; Pred. No. 6.9e-93;
 Matches 400; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 44 GCGGCTCCGGGAGCAGCAGAGAGGCGGGGAGATGTGTGCGGGGCTGCGCTCTGCG 103
 Db 2 GCGGCTCCGGGAGCAGCAGAGAGGCGGGGAGATGTGTGCGGGGCTGCGCTCTGCG 61

QY 104 GTTCCCTCCGAGCGGCGCTGATGAGCGGCACTGATTGTCCCTGGGGGCGGCGAGCGGAGCCCG 163
 Db 62 GTTCCCTCCGAGCGGCGCTGATGAGCGGCACTGATTGTCCCTGGGGGCGGCGAGCGGAGCCCG 121

QY 164 CCCGAGATGAGGCGCTGATTTAGCAAGTAAAGTAACAGAACCATGCGCTCACTTCCAA 223
 Db 122 CCCGAGATGAGGCGCTGATTTAGCAAGTAAAGTAACAGAACCATGCGCTCACTTCCAA 181
 QY 224 CACCTTTGGTGGCAGCTGATATCTGGCCATACTGTAGAGGAAGAGCAGCATG 283
 Db 182 CACCTTTGGTGGCAGCTGATATCTGGCCATACTGTAGAGGAAGAGCAGCATG 241
 QY 284 ATAGCAGTTCCTAGTTTAAACCAATATCTGTGATTCATTCTGTGATCAAGCTTGAA 343
 Db 242 ATAGCAGTTCCTAGTTTAAACCAATATCTGTGATTCATTCTGTGATCAAGCTTGAA 301
 QY 344 ACTTTTTCATCTGGGTTACTCAACTGCTTTTATGACAGATATGGGCACTAGCTG 403
 Db 302 ACTTTTTCATCTGGGTTACTCAACTGCTTTTATGACAGATATGGGCACTAGCTG 361

RESULT 13
 BF194430 547 bp mRNA linear EST 02-NOV-2000
 LOCUS BF194430
 DEFINITION 246586 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION BF194430
 VERSION BF194430.1 GI:11077799
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa

REFERENCE
 AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 1 (bases 1 to 547)
 Fahrénkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J., Vallet,J., Wise,T., Rohrer,G.A., Perlea,G., Sultana,R., Quackenbush,J. and Keefe,J.W.
 Porcine gene discovery by normalized cDNA-library sequencing and EST cluster assembly
 Mamm. Genome 13 (8), 475-478 (2002)

TITLE
 JOURNAL Contact: Smith TPL
 MEDLINE Contact: Smith TPL
 PUBMED Contact: Smith TPL
 COMMENT USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov

FEATURES
 source
 1.547
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 2P1G"
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
 library made from pooled tissue from testis, ovary,
 endometrium, hypothalamus, pituitary, and placenta."

ORIGIN
 Query Match 77.5%; Score 387.6; DB 10; Length 547;
 Best Local Similarity 93.3%; Pred. No. 3e-90;
 Matches 405; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 67 GCGGCGGAGATGTGTGCGGGGCTGCGGCTCTGCGTCCCGAGCGCGGTGAGC 126

|||||
Db 2 GGGGGGGGGGATGATGCAAGGCTGGGCTCCCTGCTCCCTTCTCGCGCGAGTGGAC 61
OY 127 GGCACGTATTTGTCCTCGGGGGGAGCGCGGACCCGCCGAGATGAGCGCTGATTAG 186
Db 62 TGCACTGATTTGTCTGTGGGGGCGAGCGCGGACCCGCTGGAATGAGCGCTCATTAG 121
OY 187 CAAGTAAATGAACAGAACCATGGCTGAGTTCCACACCTTTGGTGGCAGCTTGAT 246
Db 122 CAAGGTGAAGTAACAGAACCATGGCTGAGTTCCACACCTTTGGTGGCAGCTTGAT 181
OY 247 ATCTGGGCGCATACCTGTAGAGAGAAAGAGAGCATGATCAGAGTTCATAGTTTAAAG 306
Db 182 ATTTGGGCGCATACCTGTAGAGAGAAAGAGAGCATGATCAGAGTTCATAGCTTAAAG 241
OY 307 CCAATATCTGATTCATTACTGCTGATGATCAAGCTAGAAACTTTTTCATCTGGGTTA 366
Db 242 CCAATATCTGATTCATTACTGCTGATGATCAAGCTAGAAACTTTTTCATCTGGGTTA 301
OY 367 CCGTAACTGTTTATGACAGATATGGGCACTAGCTGACATGAATATGATGGAAGATG 426
Db 302 CCTCAACTGTTTATGACAGATATGGGCGCTGGCTGATGAATATGATGAGAGATG 361
OY 427 GATCAAGTGAAGTTTCCATAGCTATGAACTTATCAACTGAACTACAGATATGAG 486
Db 362 GATCAAGTGAAGTTTCCATAGCTATGAACTATCAACTGAACTACAGATATGAG 421
OY 487 CTACCCCTTGCACT 500
Db 422 CTCCCTCCGCACT 435

RESULT 14
CBS19152 698 bp mRNA linear EST 09-JUL-2003
LOCUS UI-M-GH0-ceh-c-14-0-UI.r1 NIH MBAP_GH0 Mus musculus cDNA clone
DEFINITION IMAGE:6839463 5', mRNA sequence.
CBS19152
ACCESSION CBS19152.1 GI:29352507
VERSION
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 698)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: ggaabs@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
The following repetitive elements were found in this cDNA
sequence: 84-130, >(CGGG)n#Simple_repeat
Seq primer: pYX-5.
Location/Qualifiers
1..698
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6839463"
/tissue_type="whole brain"
/dev_stage="1, 5, and 15 days newborn"
/lab_host="DH10B (TI phage resistant)"
/clone_id="NIH MBAP GH0"
/note="Organ: Brain; Vector: pYX-Asc; Site_1: Ecor I;

Site 2: Not I: The library was constructed according
Bonaïdo, Lemon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with Ecor I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CGAAGTGAAT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator.'

ORIGIN
Query Match 73.5%; Score 367.4; DB 14; Length 698;
Best Local Similarity 85.5%; Pred. No. 6.1e-85;
Matches 425; Conservative 0; Mismatches 61; Indels 11; Gaps 1;
OY 15 GTACGGCGGCTCGGAGAGAAATCCGAGCGGGCTCCGGAGCGA-----CAG 63
Db 58 GTACGGCGGCTCGGAGAGAGAAATCCGAGCGGGCTCCGGAGCGGAGCGAG 117
OY 64 AGAGGGCGGGGATGATGATGCGGGCGGCTCCGCTCCCTCCAGCGGCGCTG 123
Db 118 GCGGCGGGCGGGATGATGATGCGGGCGGCTCCGCTCCCTCCAGCGGCGCTG 177
OY 124 AGCGCACTGATTTGTCCTCGGGCGGAGCGGAGCCGCGGAGATGAGCGCTGAT 183
Db 178 GGCTGCACTGATTTGTGTGAGAGGGGCGGCGGCGGAGATGAGCGCTGAT 237
OY 184 TAGCAAGTAAAGTAACAGAACCATGAGCTGATTTCCACACCTTTTGGTGGAGCCTG 243
Db 238 CAGCAAGTGAACCTAATAGAACCATGAGCTGATTTCCACACCTTTGATGCTG 297
OY 244 GATATCTGGGCGCATACCTGTAGAGAGAAAGAGAGCATGATCAGAGTTCATAGTTTA 303
Db 298 GATGCTGGGCGCATACCTGTAGAGAGAAAGAGAGCATGATCAGAGTTCATAGTTTA 357
OY 304 AAGCCAAATCTGATTCATTACTGCTGATGATCAAGCTAGAAACTTTTTCATCTGGG 363
Db 358 AAGCCGATAGCGGATTTTATCTGCTGATGATCAAGCTAGAAACTTTTTCATCTGGG 417
OY 364 TTACTTCAACCTGTTTATGACAGATATGGGCACTAGCTGACATGAATATGATGAGA 423
Db 418 TTACTTCAACCTGTTTATGACAGATATGGGCGCTAGGAGCATGAATATGATGAGA 477
OY 424 ATGATCAAGTGAAGTTTCCATAGCTATGAACTTATCAACTGAACTACAGAGATAT 483
Db 478 ATGATCAAGTGAAGTTTCCATAGCTATGAACTTATCAACTGAACTACAGAGATAT 537
OY 484 CAGCTACCCCTTGCACT 500
Db 538 CAGCTCCCTCCACACT 554

RESULT 15
BC062938 5385 bp mRNA linear HTC 11-DEC-2003
LOCUS BC062938
DEFINITION Mus musculus interectin (SH3 domain protein 1A), mRNA (cDNA clone
IMAGE:6839463), containing frame-shift errors.
ACCESSION BC062938
VERSION BC062938.1 GI:38566052
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 5385)
REFERENCE Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marnusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stappleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carinini, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mallaik, J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunatane, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.N., Sodergren, E.J., Lu, X., Gibbs, R.A., Fanhey, V., Helton, E., Kettlemen, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzanski, M.I., Skelske, U., Smalhus, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932
2 (bases 1 to 5385)
Strausberg, R.
Direct Submission
Submitted (26-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: M. Bento Soares, University of Iowa
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas U. Casavant.
Web site: <http://genome.uiowa.edu>
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldo, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fishler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Sait, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: Plate: Row: Column: 0
This clone has the following problem: frame shifted.

FEATURES
source
1. 5385
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGB:4839463"
/tissue_type="Brain"
/clone_id="NIF_BMAP_GH0"
/lab_host="DH10B"
/note="Vector: pYX-ASC"

ORIGIN
Query Match 73.5%; Score 367.4; DB 11; Length 5385;
Best Local Similarity 85.5%; Pfd. No. 1e-84;
Matches 425; Conservative 0; Mismatches 61; Indels 11; Gaps 1;
QY 15 GTACGGCGGCTCGCGAGGAATCCGAGCGGCTCCGGAGACGGA-----CAG 63
DB 58 GTACGGCGGCTCGCGAGGAATCCGAGCGGCTCCGGAGCGGCGGAGGACAGGCG 117
QY 64 AGAGGCGGCGGAGATGATGTGTGCGGCGCTCGGCTCTCGTCCCTCCAGCGGCGCGTG 123
DB 118 GCGGGGCGGCGGAGATGATGTGTGCGGCGCTCGGAGCTCGGCGTTCCTCCGCGCGGCGTGCG 177
QY 124 AGCGGACCTGATTTGTCTCGGCGGCGGACCGCGGACCGCGCGGAGATGAGCGCTGAT 183
DB 178 GGCTGACCTGATTTGTGTGAGGCGGCGCGCGCGGACCGCGCGGAGATGAGCGCTGAT 237

QY 184 TAGCAAGTAAAGTAACAGAACCAATGCGCTCACTTTCCAAACACCTTTGGTGGACGCTG 243
DB 238 CAGCAAGGTGAACGTAATGAACCATGCGCTCACTTTCCAAACACCTTTGGTGGACGCTG 297
QY 244 GATATCTGGGCGCATTAAGTATAGAGGAAAGAGGAGCATGATCAGAGTTCATGTTTA 303
DB 298 GATGCTGGGCGATTAAGTATAGAGGAAAGGCGCAAGCATGACAGAGATTCCTTAGCGCTG 357
QY 304 AAGCCATATCTGATTTCTGCTGATCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 363
DB 358 AAGCCGATAGCGGAGATTTATCTGCTGATCAAGGAGAGAGCTTTTTCATCTGCGG 417
QY 364 TTACCTCAACCTGTTTACAGAGATATAGGAGCTAGTACATGAATATATGAGAGA 423
DB 418 TTACTTCAGCTGTTTACAGAGATATAGGAGCTAGTACATGAATATATGAGAGA 477
QY 424 ATGATCAAGTGAAGTTTTCATAGCTATGAACTTATCAAACTGAAGCTAAGAGATAT 483
DB 478 ATGATCAAGTGAAGTTTTCATAGCTATGAACTTATCAAACTGAAGCTAAGAGATAT 537
QY 484 CAGCTACCTCTGCACT 500
DB 538 CAGCTCCCTCCCACT 554

Search completed: July 1, 2004, 13:01:15
Job time : 2663 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2004, 08:47:35 ; Search time 388 Seconds
(without alignments)
5474.482 Million cell updates/sec

Title: US-09-720-934-1_COPY_1_500
Perfect score: 500
Sequence: 1 caaagaatccggtacgg.....taccgctaccctctgcact 500

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 3373863 segs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	500	100.0	5199	AA234570	AA234570 Human SH3
2	486	97.2	877	AAK93179	AAK93179 Human SH3
3	486	97.2	877	AAK91610	AAK91610 Human SH3
4	486	97.2	1676	AAK94611	AAK94611 Human SH3
5	486	97.2	2131	AAK94139	AAK94139 Human SH3
6	486	97.2	2131	AAH16578	AAH16578 Human SH3
7	486	97.2	3466	AA163825	AA163825 Human SH3
8	486	97.2	5195	AA234572	AA234572 Human SH3
9	486	97.2	5458	AA234571	AA234571 Human SH3
10	486	97.2	7435	AA234573	AA234573 Human SH3
11	471.4	94.3	676	AAH08146	AAH08146 Human SH3
12	444.4	88.9	3319	AAK43498	AAK43498 Human SH3
13	428	85.6	2079	AA234573	AA234573 Human SH3
14	357	71.4	5082	AA239008	AA239008 Human SH3
15	357	71.4	5738	AA239024	AA239024 Human SH3
16	345	69.0	568	AA163919	AA163919 Human SH3
17	345	69.0	568	AA231621	AA231621 Human SH3
18	345	69.0	568	AAK43814	AAK43814 Human SH3
19	241.8	48.4	3723	AA239009	AA239009 Human SH3
20	241.8	48.4	5144	AA239025	AA239025 Human SH3
21	154.4	30.6	6103	AAK52332	AAK52332 Human SH3
22	152.8	30.6	967	AA163832	AA163832 Human SH3
23	152.8	30.6	967	AA231616	AA231616 Human SH3

24	152.8	30.6	967	4	ABK43821	ABK43821 DNA encod
25	152.8	30.6	5828	6	AA147247	AA147247 Allergic
26	140.4	28.1	4625	3	AA239010	AA239010 Mouse Ese
27	140.4	28.1	6014	3	AA239026	AA239026 Mouse Ese
28	137.8	27.6	531	3	AA239026	AA239026 Mouse Ese
29	137.8	27.6	531	6	ABN72589	ABN72589 Ovarian c
30	137.8	27.6	531	8	ADA08754	ADA08754 Human ova
31	137.8	27.6	2017	3	ABN72656	ABN72656 Ovarian c
32	137.8	27.6	2017	6	ABN72656	ABN72656 Ovarian c
33	137.8	27.6	2017	8	ADA08821	ADA08821 Human ova
34	133.6	26.7	480	3	AAA69750	AAA69750 Human ova
35	133.6	26.7	480	6	ABN72644	ABN72644 Ovarian c
36	133.6	26.7	480	8	ADA08809	ADA08809 Human ova
37	133.6	26.7	3593	3	AA239011	AA239011 Mouse Ese
38	133.6	26.7	4975	3	AA239027	AA239027 Mouse Ese
39	122.8	24.6	982	6	ABQ32431	ABQ32431 Oligonuci
40	122.8	24.6	982	6	ABQ32430	ABQ32430 Oligonuci
41	122.4	24.5	462	8	ACH34516	ACH34516 Human end
42	113.2	22.6	4447	4	AA239055	AA239055 DNA encod
43	109	21.8	292	4	AA122800	AA122800 Probe #12
44	109	21.8	292	4	ABA67887	ABA67887 Human foe
45	109	21.8	292	4	AA148103	AA148103 Probe #16

ALIGNMENTS

RESULT 1	AA234570	standard; cDNA; 5199 BP.
ID	AA234570	
AC	AA234570;	
DT	01-FEB-2000	(first entry)
DE	Human SH3D1A cDNA.	
XX	SH3D1A gene; human; Down's syndrome; leukaemia; cancer;	
KW	megakaryocytic abnormality; myeloproliferative disorder;	
KW	platelet disorder; neural disorder; thrombocytopenia;	
KW	haematopoietic disorder; cognitive dysfunction; microcephaly;	
KW	lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.	
OS	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	CDS	208..3642
FT		/*tag= a
XX	MO9953062-A2.	
XX	21-OCT-1999.	
XX	16-APR-1999;	99WO-US008371.
XX	16-APR-1999;	98US-0082007P.
XX	(CEDA-) CEDARS SINAI HEALTH SYSTEM.	
XX	Korenberg JR, Chen X;	
XX	WPI: 1999-633829/54.	
XX	P-PSDB; AA132154.	
XX	Nucleic acid from the human SH3D1A gene and its products, useful for the	
XX	diagnosis and treatment of myeloproliferative disorders and leukemia.	
XX	Claim 2; Fig 5; 99pp; English.	
XX	This is the nucleotide sequence of full-length cDNA corresponding to a	
XX	novel human SH3 gene, termed the SH3D1A gene, that contributes to the	
XX	development of platelets and the pathogenesis of leukemias, both in	
XX	general and in particular those involving the megakaryocytic lineage. The	

SH3D1A gene maps to the small candidate region for low platelets on chromosome 21. Sequencing of 5 different sizes of cDNA clone (see AA234570-74) suggests that at least 3 isoforms exist. The invention provides methods for the diagnosis and treatment of megakaryocytic abnormalities, myeloproliferative disorder, platelet disorder, acute leukaemia, neural disorders, thrombocytopenia, platelet disorder on chromosome 21, low platelets in deletion for 21, association of gains in chromosome 21 with leukaemias, neural abnormalities, dysfunction of gains in disorders including brain malformations and corresponding cognitive and dysfunctions, microcephaly, lissencephaly, and colpocephaly. Methods are also provided for: suppressing cells unable to regulate themselves; screening for a somatic alteration in the SH3D1A gene; monitoring the progress and adequacy of a treatment; monitoring tumour risk progress or megakaryocytic abnormality, myeloproliferative disorder, haematopoietic disorder, platelet disorder or leukaemia; and treatment of a subject (including a prenatal subject) having megakaryocytic abnormality, myeloproliferative disorder, platelet disorder, leukaemia or neural disorder using a nucleic acid that expresses SH3D1A or its antisense nucleic acid

Sequence 5199 BP; 1615 A; 1152 C; 1214 G; 1218 T; 0 U; 0 Other;

Query Match 100.0%; Score 500; DB 2; Length 5199;

Best Local Similarity 100.0%; Pred. No. 3e-142; Mismatches 0; Gaps 0;

Matches 500; Conservative 0; Indels 0; Gaps 0;

```

QY 1 CAAAGAGATTCGCGGTACGCGCTCGCGAGAGAAATCCCGAGCGGGCTCGGGGACGGA 60
DB 1 CAAAGAGATTCGCGGTACGCGCTCGCGAGAGAAATCCCGAGCGGGCTCGGGGACGGA 60
QY 61 CAGAGAGCGGGCGGGGATGCTGCGGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGG 120
DB 61 CAGAGAGCGGGCGGGGATGCTGCGGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGG 120
QY 121 GCGAGCGGCGGATGCTGCTGCGGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGG 180
DB 121 GCGAGCGGCGGATGCTGCTGCGGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGG 180
QY 181 GATTAGCAAGGTAAAGTAACAGAACCATGCTCACTTTCACACCTTTTGGTGGCAG 240
DB 181 GATTAGCAAGGTAAAGTAACAGAACCATGCTCACTTTCACACCTTTTGGTGGCAG 240
QY 241 CTGAGATTCGCGGCTTAACCTGTAGAGAAAGCAAGCATGATCAGAGCTTCATAGT 300
DB 241 CTGAGATTCGCGGCTTAACCTGTAGAGAAAGCAAGCATGATCAGAGCTTCATAGT 300
QY 301 TTAAGCCAAATCTGAGATTCATTACTGATGATCAAGTAACTTTTTCATCT 360
DB 301 TTAAGCCAAATCTGAGATTCATTACTGATGATCAAGTAACTTTTTCATCT 360
QY 361 GGGTTAAGCTCAACCTGTTTAAACAGATATGGGCACTAGTGAATATATATGA 420
DB 361 GGGTTAAGCTCAACCTGTTTAAACAGATATGGGCACTAGTGAATATATATGA 420
QY 421 AATAATGATCAAGTGAAGTTTCCATAGCTATGAATCTTATCAAACTGAAGCTACAGGA 480
DB 421 AATAATGATCAAGTGAAGTTTCCATAGCTATGAATCTTATCAAACTGAAGCTACAGGA 480
QY 481 TATCAGCTACCTCTGCACT 500
DB 481 TATCAGCTACCTCTGCACT 500

```

RESULT 2

AAK93179

AAK93179 standard; cDNA; 877 BP.

AAK93179;

06-NOV-2001 (first entry)

Human cDNA clone representative sequence, SEQ ID NO: 1639.

XX

KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

OS Homo sapiens.

PN EPI130094-A2.

PD 05-SEP-2001.

PF 07-JUL-2000; 2000EP-00114089.

PR 08-JUL-1999; 99JP-00194486.

PR 11-JAN-2000; 2000JP-00118774.

PR 02-MAY-2000; 2000JP-00183765.

XX (HELI-) HELIX RES INST.

PI Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

DR WPI; 2001-524255/58.

PT 830 Primers useful for synthesizing full length cDNA clones and their use

PS in genetic manipulation.

XX Example 11; SEQ ID NO 1639; 1380bp + Sequence Listing; English.

XX The invention relates to primers for synthesizing full length cDNA

CC clones. 830 cDNA molecules encoding a human protein have been isolated

CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have

CC been determined. Primers for synthesizing the full length cDNA are useful

CC for clarifying the function of the protein encoded by the cDNA. The full

CC length clones were obtained by construction of full length enriched cDNA

CC libraries that were synthesised by the oligo-capping method. The primers

CC enable the production of the full length cDNA easily without any special

CC methods. The present sequence was used as the representative sequence

CC from a human clone which was used in homology searches to identify the

CC clone. Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in CD-ROM format directly from

CC EPO

XX Sequence 877 BP; 208 A; 228 C; 223 G; 213 T; 0 U; 5 Other;

XX Query Match 97.2%; Score 486; DB 4; Length 877;

XX Best Local Similarity 100.0%; Pred. No. 2.4e-138; Mismatches 0; Gaps 0;

XX Matches 486; Conservative 0; Indels 0; Gaps 0;

QY 15 GTACGGCGGCTCGCGAGAGAAATCCCGAGCGGGCTCGGGACGAGACAGAGGCGGGCG 74

DB 12 GTACGGCGGCTCGCGAGAGAAATCCCGAGCGGGCTCGGGACGAGACAGAGGCGGGCG 71

QY 75 GGGATGGTGTGCGGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGG 134

DB 72 GGGATGGTGTGCGGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGG 131

QY 135 TTTGTCCTCGGGGCGGACGCGGACCCGCGGAGATGAGCGTGCATTAAGCAAGTAA 194

DB 132 TTTGTCCTCGGGGCGGACGCGGACCCGCGGAGATGAGCGTGCATTAAGCAAGTAA 191

QY 195 AAGTAAACAGAACCATGCTCACTTTTTCACACCTTTTGGTGGACGCTGGATATCTGGCC 254

DB 192 AAGTAAACAGAACCATGCTCACTTTTTCACACCTTTTGGTGGACGCTGGATATCTGGCC 251

QY 255 CATTAAGTGAAGAGAGAGAGAGATGATGAGCATGAGCTTCACTTTTAAAGCCAAATATC 314

DB 252 CATTAAGTGAAGAGAGAGAGATGATGAGCATGAGCTTCACTTTTAAAGCCAAATATC 311

QY 315 TGGATTCATTACTGATGATCAAGCTAGAAACTTTTTCATCTGGGTTAAGCTCAACG 374

DB 312 TGGATTCATTACTGATGATCAAGCTAGAAACTTTTTCATCTGGGTTAAGCTCAACG 371

QY 375 TGTTTTACACAGATATGAGGCACTAGCTGACATGAATATGATGAAGATGATCAAGT 434

DB 372 TGTTTTACACAGATATGAGGCACTAGCTGACATGAATATGATGAAGATGATCAAGT 431

QY 435 GGAGTTTCCATAGCTATGAAACTTATCAAACTGAGCTACAGATATCAGTACCTC 494
DB 432 GGAGTTTCCATAGCTATGAAACTTATCAAACTGAGCTACAGATATCAGTACCTC 491
QY 495 TGCACT 500
DB 492 TGCACT 497

RESULT 3
AAK91610
ID AAK91610 standard; cDNA; 877 BP.
AAK91610;
AC 06-NOV-2001 (first entry)
XX
DE Human cDNA 5'-end sequence, SEQ ID NO: 70.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX OS Homo sapiens.
XX PN EP1130094-A2.
XX PD 05-SEP-2001.
XX PE 07-JUL-2000; 2000BP-00114089.
XX PR 08-JUL-1999; 99JP-00194486.
XX PR 11-JAN-2000; 2000JP-00118774.
XX PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
XX DR 830 Primers useful for synthesizing full length cDNA clones and their use
XX PT in genetic manipulation.
XX PS Claim 2; SEQ ID NO 70; 1380BP + Sequence Listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesized by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is the nucleotide sequence of the 5'-end of
CC a cDNA provided in the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in CD-
CC ROM format directly from EPO
XX

SEQ Sequence 877 BP; 208 A; 228 C; 223 G; 213 T; 0 U; 5 Other;
Query Match 97.2%; Score 486; DB 4; Length 877;
Best Local Similarity 100.0%; Pred. No. 2.4e-138;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GTACGGCGGCTGCGGAGGAATCCGAGCGGCTCGGGGACGAGCAGAGAGGCGGCG 74
DB 12 GTACGGCGGCTGCGGAGGAATCCGAGCGGCTCGGGGACGAGCAGAGAGGCGGCG 71
QY 75 GGAGTGTGTGCGGGGCTCGGCTCCTGCTCCCTCCAGCGCGGTGAGGCGGCACTGA 134
DB 72 GGAGTGTGTGCGGGGCTCGGCTCCTGCTCCCTCCAGCGCGGTGAGGCGGCACTGA 131

QY 135 TTTGTCCCTGGGGGCGGAGCGGACCCGCCGAGATGAGGCGTGCATTAGCAAGGTAA 194
DB 132 TTTGTCCCTGGGGGCGGAGCGGACCCGCCGAGATGAGGCGTGCATTAGCAAGGTAA 191
QY 195 AAGTACAGAACCATGAGCTCAGTTTCCAGACCTTTTGGTGGAGCTGATATCTGGGC 254
DB 192 AAGTACAGAACCATGAGCTCAGTTTCCAGACCTTTTGGTGGAGCTGATATCTGGGC 251
QY 255 CATTAAGTGTGAGGAGAAAGAGCGAAGCATGATCAGCAGTTCCATGTTTAAAGCCAAATTC 314
DB 252 CATTAAGTGTGAGGAGAAAGAGCGAAGCATGATCAGCAGTTCCATGTTTAAAGCCAAATTC 311
QY 315 TGGAATCATACGTCGTATCAAGCTAGAACTTTTTCATCTGGGTTACCTCAACC 374
DB 312 TGGAATCATACGTCGTATCAAGCTAGAACTTTTTCATCTGGGTTACCTCAACC 371
QY 375 TGTTTAGCAGATATGAGGACCTAGCTGACATGATATGATGAGAAATGATCAAGT 434
DB 372 TGTTTAGCAGATATGAGGACCTAGCTGACATGATATGATGAGAAATGATCAAGT 431
QY 435 GGAGTTTCCATAGCTATGAAACTTATCAAACTGAGCTACAGATATCAGTACCTC 494
DB 432 GGAGTTTCCATAGCTATGAAACTTATCAAACTGAGCTACAGATATCAGTACCTC 491
QY 495 TGCACT 500
DB 492 TGCACT 497

RESULT 4
AAK94611
ID AAK94611 standard; cDNA; 1676 BP.
AAK94611;
AC 07-NOV-2001 (first entry)
XX
DE Human cDNA, SEQ ID NO: 3565.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX OS Homo sapiens.
XX PN EP1130094-A2.
XX PD 05-SEP-2001.
XX PE 07-JUL-2000; 2000BP-00114089.
XX PR 08-JUL-1999; 99JP-00194486.
XX PR 11-JAN-2000; 2000JP-00118774.
XX PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
XX DR P-PSDB; AAM93676.
XX PT 830 Primers useful for synthesizing full length cDNA clones and their use
XX PT in genetic manipulation.
XX PS Disclosure; SEQ ID NO 3565; 1380BP + Sequence Listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesized by the oligo-capping method. The primers

AAH16578;
XX 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:15658.
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-00116126.
XX 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX
XX
PS Claim 8; SEQ ID NO 15658; 2537bp + Sequence listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH16742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 2131 BP; 691 A; 448 C; 572 G; 420 T; 0 U; 0 Other;
Query Match 97.2%; Score 486; DB 4; Length 2131;
Best Local Similarity 100.0%; Pred. No. 3.8e-138;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 TTGTCCCTGGGGGCGGAGCGGAGCCCGCGGAGTAGAGGCTGATTAGCAAGGTAA 194
DB 274 TTTGTCCCTGGGGGCGGAGCGGAGCCCGCGGAGTAGAGGCTGATTAGCAAGGTAA 333
QY 195 AAGTAAACAGAACCATGGCTCAGTTTCCACACCTTTTGTGGAGCCTGTGATATCTGGGC 254
DB 334 AAGTAAACAGAACCATGGCTCAGTTTCCACACCTTTTGTGGAGCCTGTGATATCTGGGC 393
QY 255 CATTAAGTGTAGAGAAAGAGCGAAGCATGATCAGCAGTTCCATGTTTAAAGCCATATC 314
DB 394 CATTAAGTGTAGAGAAAGAGCGAAGCATGATCAGCAGTTCCATGTTTAAAGCCATATC 453
QY 315 TGATTTCATTAATCAGTGTATCAAGCTGAATCTTTTTCATCTGGGTACTCAACC 374
DB 454 TGATTTCATTAATCAGTGTATCAAGCTGAATCTTTTTCATCTGGGTACTCAACC 513
QY 375 TGTTTAGCAGATATAGGCACTAGCTGCATGATGATATGATGAAATGATCAAGT 434
DB 514 TGTTTAGCAGATATAGGCACTAGCTGCATGATGATATGATGAAATGATCAAGT 573
QY 435 GGAAGTTTTCATAGCTATGAAATTTATCAAACTGAAGCTACAGGATATCAGTACCTC 494
DB 574 GGAAGTTTTCATAGCTATGAAATTTATCAAACTGAAGCTACAGGATATCAGTACCTC 633
QY 495 TGCACT 500
DB 634 TGCACT 639
RESULT 7
AA163825
ID AA163825 standard; cDNA; 3466 BP.
XX AA163825;
DT 22-OCT-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 33.
DB
XX
KW Human; antiarthritic; antirheumatic; antiproliferative; vasotropic;
KW cerebroprotective; nootropic; neuroprotective; antibacterial; virocidic;
KW fungicide; ophthalmological; cytostatic; immunosuppressive; nootropic;
KW neuroprotective; antiallergic; hepatotropic; antidiabetic;
KW antiinflammatory; antitumor; vulnary; anticonvulsant; antibacterial;
KW antiparasitic; cardiant; gene therapy; cancer; immune disorder;
KW cardiovascular disorder; neurological disease; infection; human; ss.
XX
OS Homo sapiens.
XX
PN WO200155308-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001309.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0188874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.

CC form part of the printed specification, but was obtained in electronic
 CC format directly from MIP0 at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 3466 BP; 1056 A; 829 C; 861 G; 710 T; 0 U; 10 Other;

Query Match 97.2%; Score 486; DB 4; Length 3466;

Best Local Similarity 100.0%; Pred. No. 4.8e-138; Indels 0; Gaps 0;

Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GTACGGCGGCTCGGAGAGAAATCCGAGCGGCTCCGGACGACAGAGAGCGGCG 74
 DB 84 GTACGGCGGCTCGGAGAGAAATCCGAGCGGCTCCGGACGACAGAGAGCGGCG 143
 QY 75 GGGATGCTGCGGGGCTGCGGCTCTGCTCCCTCCAGCGGCGCTGAGCGGACTGA 134
 DB 144 GGGATGCTGCTGCGGGGCTGCGGCTCTGCTCCCTCCAGCGGCGCTGAGCGGACTGA 203
 QY 135 TTTGTCCTCGGGGCGGACGCGGACCGCGGAGATGAGAGCGGCTGAGCAAGGTAA 194
 DB 204 TTTGTCCTCGGGGCGGACGCGGACCGCGGAGATGAGAGCGGCTGAGCAAGGTAA 263
 QY 195 AAGTAACGAAACCATGCTCAAGTTTCCAAACCTTTTGTGCGACCTGATATCTGCGC 254
 DB 264 AAGTAACGAAACCATGCTCAAGTTTCCAAACCTTTTGTGCGACCTGATATCTGCGC 323
 QY 255 CATTAAGTATGAGGAAAGAGGAGCATGATGAGATGATGATGATGATGATGATGATG 314
 DB 324 CATTAAGTATGAGGAAAGAGGAGCATGATGAGATGATGATGATGATGATGATGATG 383
 QY 315 TGGATTCATTAAGTATGAGGAAAGAGGAGCATGATGAGATGATGATGATGATGATG 374
 DB 384 TGGATTCATTAAGTATGAGGAAAGAGGAGCATGATGAGATGATGATGATGATGATG 443
 QY 375 TGTTTTACGACAGATATGAGGACCTAGTGCATGATGATGATGATGATGATGATGATG 434
 DB 444 TGTTTTACGACAGATATGAGGACCTAGTGCATGATGATGATGATGATGATGATGATG 503
 QY 435 GGAGTTTTCATGCTATGAAACTTATCAAACTGAAAGTACAGAGATATCACTACCTC 494
 DB 504 GGAGTTTTCATGCTATGAAACTTATCAAACTGAAAGTACAGAGATATCACTACCTC 563
 QY 495 TGGACT 500
 DB 564 TGGACT 569

RESULT 8
 AA234572 standard: cDNA, 5195 BP.
 ID AA234572
 XX
 AC AA234572;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human SH3D1A cDNA clone 11.
 XX
 KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 239..3886
 ET /*tag= a
 XX
 PN WO9953062-A2.
 XX
 PD 21-OCT-1999.
 XX
 PF 16-APR-1999; 99WO-US008371.

XX
 PR 16-APR-1999; 98US-0082007P.
 XX
 PA (CEDA-) CEDARS SIGNAL HEALTH SYSTEM.
 XX
 PI Korenberg JR, Chen X;
 XX
 DR WPI, 1999-633829/54.
 DR P-PSDI, AAY32156.
 XX
 PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 DB diagnosis and treatment of myeloproliferative disorders and leukemia.
 XX
 PS Claim 2; Fig 10; 99pp; English.
 XX
 CC This is the nucleotide sequence of full-length cDNA (clone 11)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC platelet disorder on chromosome 21, low platelets in deletion for 21,
 CC association of gains in chromosome 21 with leukemias, neural
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukaemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 XX
 SQ Sequence 5195 BP; 1592 A; 1148 C; 1237 G; 1218 T; 0 U; 0 Other;

Query Match 97.2%; Score 486; DB 2; Length 5195;
 Best Local Similarity 100.0%; Pred. No. 5.9e-138;
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GTACGGCGGCTCGGAGAGAAATCCGAGCGGCTCCGGACGACAGAGAGCGGCG 74
 DB 46 GTACGGCGGCTCGGAGAGAAATCCGAGCGGCTCCGGACGACAGAGAGCGGCG 105
 QY 75 GGGATGCTGCGGGGCTGCGGCTCTGCTCCCTCCAGCGGCGCTGAGCGGACTGA 134
 DB 106 GGGATGCTGCGGGGCTGCGGCTCTGCTCCCTCCAGCGGCGCTGAGCGGACTGA 165
 QY 135 TTTGTCCTCGGGGCGGACGCGGACCGCGGAGATGAGGCGTGAATTAGCAAGGTAA 194
 DB 166 TTTGTCCTCGGGGCGGACGCGGACCGCGGAGATGAGGCGTGAATTAGCAAGGTAA 225
 QY 195 AAGTAACGAAACCATGCTCAAGTTTCCAAACCTTTTGTGCGACCTGATATCTGCGC 254
 DB 226 AAGTAACGAAACCATGCTCAAGTTTCCAAACCTTTTGTGCGACCTGATATCTGCGC 285
 QY 255 CATTAAGTATGAGGAAAGAGGAGCATGATGAGATGATGATGATGATGATGATGATG 314
 DB 286 CATTAAGTATGAGGAAAGAGGAGCATGATGAGATGATGATGATGATGATGATGATG 345
 QY 315 TGGATTCATTAAGTATGAGGAAAGAGGAGCATGATGAGATGATGATGATGATGATG 374
 DB 346 TGGATTCATTAAGTATGAGGAAAGAGGAGCATGATGAGATGATGATGATGATGATG 405
 QY 375 TGTTTTACGACAGATATGAGGACCTAGTGCATGATGATGATGATGATGATGATGATG 434
 DB 406 TGTTTTACGACAGATATGAGGACCTAGTGCATGATGATGATGATGATGATGATGATG 465

QY 435 GGAGTTTTCATGCTATGAACCTTATCAACTGAGCTCAAGAGATATCAGTACCCCTC 494
 DB 466 GGAGTTTTCATGCTATGAACCTTATCAACTGAGCTCAAGAGATATCAGTACCCCTC 525
 QY 495 TGCACCT 500
 DB 526 TGCACCT 531

RESULT 9
 AA234571
 ID AA234571 standard; cDNA; 5458 BP.
 XX AA234571;
 AC
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human SH3D1A cDNA clone 21.
 XX
 XX SH3D1A gene; human; Down's syndrome; leukemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haemopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 267..3929
 FT /*tag= a
 XX
 PN W09953062-A2.
 XX
 PD 21-OCT-1999.
 XX
 XX 16-APR-1999; 99WO-US008371.
 PF
 XX
 PR 16-APR-1998; 98US-0082007P.
 XX
 PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 XX
 PI Korenberg JR, Chen X;
 XX
 DR WPI; 1999-633829/54.
 DR P-PSDB; AAY32155.
 XX
 PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 XX
 PS Claim 2; Fig 8; 99P; English.
 XX
 CC This is the nucleotide sequence of full-length cDNA (clone 21)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukemia, neural disorders, thrombocytopenia,
 CC association of gains in chromosome 21, low platelets in deletion for 21,
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukemia or neural disorder using a nucleic acid that

CC expresses SH3D1A or its antisense nucleic acid
 XX
 SQ Sequence 5458 BP; 1671 A; 1201 C; 1307 G; 1279 T; 0 U; 0 Other;
 Query Match 97.2%; Score 486; DB 2; Length 5458;
 Best Local Similarity 100.0%; Pred. No. 6,1e-138;
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GTACGGCGGCTGCGAGAGAAATCCCGAGCGGCTCCGGACGCGACAGAGAGCGGGCG 74
 DB 74 GTACGGCGGCTGCGAGAGAAATCCCGAGCGGCTCCGGACGCGAGAGAGCGGGCG 133
 QY 75 GGGATGGTGGCGGGGCTGGCGCTCCTCGCTCCCTCCAGCGGGCTGAGCGGCACTGA 134
 DB 134 GGGATGGTGGCGGGGCTGGCGCTCCTCGCTCCCTCCAGCGGGCTGAGCGGCACTGA 193
 QY 135 TTTGTCCTGGGCGGCGAGCGGACCGCGCGAGATGAGCGTCGATTAGCAAGGTAA 194
 DB 194 TTTGTCCTGGGCGGCGAGCGGACCGCGCGAGATGAGCGTCGATTAGCAAGGTAA 253
 QY 195 AAGTAAACAGAACCATGGCTGAGTTTCCACACCTTTTGTGGGAGCTGAGATATCGGGC 254
 DB 254 AAGTAAACAGAACCATGGCTGAGTTTCCACACCTTTTGTGGGAGCTGAGATATCGGGC 313
 QY 255 CATACCTGTAGAGAAAGCGAAGCATGATGACAGTTCCATTGTTTAAAGCAATATC 314
 DB 314 CATACCTGTAGAGAAAGCGAAGCATGATGACAGTTCCATTGTTTAAAGCAATATC 373
 QY 315 TGGATTCAATTACTGTGATCACTAGTAAACCTTTTTCATCTGGTTACTCAACG 374
 DB 374 TGGATTCAATTACTGTGATCACTAGTAAACCTTTTTCATCTGGTTACTCAACG 433
 QY 375 TGTTTAGACAGATATGGGCACTAGCTGACATGAATATGATGAAAGATGATCAAGT 434
 DB 434 TGTTTAGACAGATATGGGCACTAGCTGACATGAATATGATGAAAGATGATCAAGT 493
 QY 435 GGAGTTTTCATGCTATGAACCTTATCAACTGAGCTCAAGAGATATCAGTACCCCTC 494
 DB 494 GGAGTTTTCATGCTATGAACCTTATCAACTGAGCTCAAGAGATATCAGTACCCCTC 553
 QY 495 TGCACCT 500
 DB 554 TGCACCT 559

RESULT 10
 AAS84763
 ID AAS84763 standard; cDNA; 7435 BP.
 XX
 XX AAS84763;
 AC
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #20567.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.
XX P-PSDB; ABG20576.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 20567; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (II) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 7435 BP; 2190 A; 1690 C; 1829 G; 1726 T; 0 U; 0 Other;
SQ
Query Match 96.9%; Score 484.4; DB 5; Length 7435;
Best Local Similarity 99.8%; Pred. No. 2,2e-137;
Matches 485; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 15 GTACGGCGGCTCGCGAGGAAGATCCGAGCGGGCTCCGGAGCGAGACAGAGCGGCG 74
DB GTACGGCGGCTCGCGAGGAAGATCCGAGCGGGCTCCGGAGCGAGACAGAGCGGCG 129
XX
XX 70 GTACGGCGGCTCGCGAGGAAGATCCGAGCGGGCTCCGGAGCGAGACAGAGCGGCG 129
XX
XX 75 GGAGTGTGTGCGGGGCTGCGGCTCTGCTCCCTCCAGCGGCGCGTGAAGCGGACGTA 134
DB GGAGTGTGTGCGGGGCTGCGGCTCTGCTCCCTCCAGCGGCGCGTGAAGCGGACGTA 189
XX
XX 130 GGAGTGTGTGCGGGGCTGCGGCTCTGCTCCCTCCAGCGGCGCGTGAAGCGGACGTA 189
DB GGAGTGTGTGCGGGGCTGCGGCTCTGCTCCCTCCAGCGGCGCGTGAAGCGGACGTA 249
XX
XX 135 TTTGTCCTGGGGCGGCGAGCGGCGCGCGCGAGATGAGGCGCTGATTAGCAAGTAA 194
DB TTTGTCCTGGGGCGGCGAGCGGCGCGCGCGAGATGAGGCGCTGATTAGCAAGTAA 249
XX
XX 190 TTTGTCCTGGGGCGGCGAGCGGCGCGCGCGAGATGAGGCGCTGATTAGCAAGTAA 249
DB TTTGTCCTGGGGCGGCGAGCGGCGCGCGCGAGATGAGGCGCTGATTAGCAAGTAA 254
XX
XX 195 AAGTAACAGAACATGGCTCACTTTTCCAAACCTTTTGGTGGAGCTTGATATCTGGGC 254
DB AAGTAACAGAACATGGCTCACTTTTCCAAACCTTTTGGTGGAGCTTGATATCTGGGC 309
XX
XX 250 AAGTAACAGAACATGGCTCACTTTTCCAAACCTTTTGGTGGAGCTTGATATCTGGGC 309
DB AAGTAACAGAACATGGCTCACTTTTCCAAACCTTTTGGTGGAGCTTGATATCTGGGC 314
XX
XX 255 CATACCTGAGGAAAGGAGCAAGCATGATGAGAGTTCCATGTTAAACCAATATC 314
DB CATACCTGAGGAAAGGAGCAAGCATGATGAGAGTTCCATGTTAAACCAATATC 369
XX
XX 310 CATACCTGAGGAAAGGAGCAAGCATGATGAGAGTTCCATGTTAAACCAATATC 369
DB CATACCTGAGGAAAGGAGCAAGCATGATGAGAGTTCCATGTTAAACCAATATC 374
XX
XX 315 TGGATTCTTACTGGTGAATCAAGTAAACTTTTTTTTCAATCTGGTTACTCAAC 374
DB TGGATTCTTACTGGTGAATCAAGTAAACTTTTTTTTCAATCTGGTTACTCAAC 429
XX
XX 370 TGGATTCTTACTGGTGAATCAAGTAAACTTTTTTTTCAATCTGGTTACTCAAC 429
DB TGGATTCTTACTGGTGAATCAAGTAAACTTTTTTTTCAATCTGGTTACTCAAC 434
XX
XX 375 TGGTTTGAACAGATATGGGCACTAGCTGACATGATTAATGATGAAGATGATCACT 434
DB TGGTTTGAACAGATATGGGCACTAGCTGACATGATTAATGATGAAGATGATCACT 489
XX
XX 430 TGGTTTGAACAGATATGGGCACTAGCTGACATGATTAATGATGAAGATGATCACT 489
DB TGGTTTGAACAGATATGGGCACTAGCTGACATGATTAATGATGAAGATGATCACT 494
XX
XX 435 GGAGTTTCCATAGCTATGAAACTTATCAACTGAAGTACAGAGTATCACTCACTCC 494
DB GGAGTTTCCATAGCTATGAAACTTATCAACTGAAGTACAGAGTATCACTCACTCC 549
XX
XX 490 GGAGTTTCCATAGCTATGAAACTTATCAACTGAAGTACAGAGTATCACTCACTCC 549
DB GGAGTTTCCATAGCTATGAAACTTATCAACTGAAGTACAGAGTATCACTCACTCC 555
XX
XX 550 TGCACCT 555
DB TGCACCT 555

RESULT 11
AAH08146
ID AAH08146 standard; cDNA; 676 BP.
XX
XX AAH08146;
AC
XX 26-JUN-2001 (first entry)
DT
XX Human cDNA clone (5'-primer) SEQ ID NO:4981.
DE
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
KW
XX Homo sapiens.
OS
XX EP1074617-A2.
PN
XX 07-FEB-2001.
PD
XX 28-JUN-2000; 2000EP-00116126.
PF
XX 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PS 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
PA
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI: 2001-318749/34.
DR
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
XX Claim 1; SEQ ID NO 4981; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH01166 to AAH13628 and
CC AAH13633 to AAH16742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH11629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
XX Sequence 676 BP; 167 A; 141 C; 228 G; 137 T; 0 U; 3 Other;
SQ
Query Match 94.3%; Score 471.4; DB 4; Length 676;
Best Local Similarity 99.4%; Pred. No. 6.3e-134;
Matches 483; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
XX
XX 15 GTACGGCGGCTCGCGAGGAAGATCCGAGCGGGCTCCGGAGCGAGACAGAGCGGCG 74
DB GTACGGCGGCTCGCGAGGAAGATCCGAGCGGGCTCCGGAGCGAGACAGAGCGGCG 74

DB 154 GTACGGGCGCTCCGACGAGAAATCCCGACGGGCTCCGGGACGACAGAGCGGGCG 213
QY 75 GGGATGGTGTGGGGGCTGGGCTCTCCCTGGCTCCACGGGGCTGAGCGGCACACTGA 134
DB 214 GGGATGGTGTGGGGGCTGGGCTCTCCCTGGCTCCACGGGGCTGAGCGGCACACTGA 273
QY 135 TTGTCCCTGGGGCGGCGAGCGCGGACCCGGCCGAGATGAGGGGTCGATTAGCAAGTAA 194
DB 274 TTGTCCCTGGGGCGGCGAGCGCGGACCCGGCCGAGATGAGGGGTCGATTAGCAAGTAA 333
QY 195 AAGTAACAGAACCATGGCTAGTTTCCAAACCTTTTGGTGGACGCTGGATATCTGGGC 254
DB 334 AAGTAACAGAACCATGGCTAGTTTCCAAACCTTTTGGTGGACGCTGGATATCTGGGC 393
QY 255 CATTAAGTGTAGAGAAAGAGCGAAGCATGATCCAGAGTTCATAGTTTAAGCCATATC 314
DB 394 CATTAAGTGTAGAGAAAGAGCGAAGCATGATCCAGAGTTCATAGTTTAAGCCATATC 453
QY 315 TGGATTCATTACTGGTGATCAAGCTAGAAACCTTTTTCAAATCTGGTTACCTCAAC 374
DB 454 TGGATTCATTACTGGTGATCAAGCTAGAAACCTTTTTCAAATCTGGTTACCTCAAC 513
QY 375 TGTTTAGCACACATATGGGCTAGTGTGACATGATGATGAGAGATGATCAAGT 434
DB 514 TGTTTAGCACACATATGGGCTAGTGTGACATGATGATGAGAGATGATG-TCAGT 572
QY 435 GGAGTTTTCATAGCTATGTAACCTTATCAAGTGAAGCTACAGGATATGAGTACCTTC 494
DB 573 TGAATTTTCATATGCTATGTAACCTTATCAAGTGAAGCTACAGGATATGAGTACCTTC 632
QY 495 TGCACCT 500
DB 633 TGCACCT 638
RESULT 12
ABK43498
ID ABK43498 standard; cDNA; 3319 bp.
XX
AC ABK43498;
XX
DT 05-JUN-2002 (first entry)
XX
DE DNA encoding novel central nervous system protein #78.
XX
KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200155318-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US001332.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234980P.
PR 25-SEP-2000; 2000US-0234981P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.

CC	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241787P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	20-OCT-2000;	2000US-0241809P.
PR	20-OCT-2000;	2000US-0241826P.
PR	01-NOV-2000;	2000US-0246471P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0246521P.
PR	08-NOV-2000;	2000US-0246524P.
PR	08-NOV-2000;	2000US-0246525P.
PR	08-NOV-2000;	2000US-0246526P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246609P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249264P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	03-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0251989P.
PR	06-DEC-2000;	2000US-0251990P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251983P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259678P.
XX		
PA	(HUMA-)	HUMAN GENOME SCI INC.
XX		
PI	Rosen CA,	Barash SC, Ruben SM,
DR	WPI;	2001-581633/65.
XX	P-PSDB;	AAU87168.
XX		
PT	New isolated	nucleic acid encoding a protein for diagnosing, preventing,
PT	treating or	ameliorating medical conditions and used as food additives or
PT	preservatives.	
XX		
PS	Claim 1;	SEQ ID NO 88; 837pp; English.
XX		
CC	The invention	describes an isolated nucleic acid molecule (II) encoding a
CC	novel central	nervous system protein. (i) and polypeptides (III) encoded
CC	by (I), are	used to treat a medical conditions and in diagnosis of a
CC	pathological	condition. Disorders which are diagnosed or treated include
CC	autoimmune	diseases e.g. rheumatoid arthritis, hyperproliferative

CC disorders e.g. neoplasms of the breast or liver cardiovascular disorders
CC disorders e.g. cerebrovascular disorders e.g. cerebral ischemia,
CC e.g. cardiac arrest, nervous system disorders e.g. Alzheimer's disease and
CC angiogenesis, nervous system disorders e.g. infections caused by bacteria, viruses
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC adenocarcinomas and irritiable bowel syndrome, reproductive system
CC disorders e.g. testicular feminization, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polyphenols can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polyphenols can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match	88.9%	Score 444.4;	DB 4;	Length 3319;
Best Local Similarity	99.8%;	Pred. No. 2.7e-125;		
Matches 445;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0

OY	5	GAAGGACAAGAGCGGGCGGGGATGAGTGTGGCGGGGCTGCGGCTCTGCGCTCCAG	114
Db	6	GAAGGACAAGAGCGGGCGGGGATGAGTGTGGCGGGCTGCGGCTCTGCGCTCCAG	65
OY	115	CGGCGGTGAGCGGCATGATTTGCTCCCTGCGGCGCGCGGACCCGCGGAGTGA	174
Db	66	CGGCGGTGAGCGGCATGATTTGCTCCCTGCGGCGCGCGGACCCGCGGAGTGA	125
OY	175	GGGCTGATTAGCAAGGTAAAGTAAAGAACATAGGCTGAGTTTCCAAACCTTTGGT	234
Db	126	GGGCTGATTAGCAAGGTAAAGTAAAGAACATAGGCTGAGTTTCCAAACCTTTGGT	185
OY	235	GGCAGCTGGATATCTGGGSCATACCTAGAGGAAAGGAGACATGATCAGCAATTC	294
Db	186	GGCAGCTGGATATCTGGGSCATACCTAGAGGAAAGGAGACATGATCAGCAATTC	245
OY	295	CATAGTTTAAAGCAATATCTGAGTTCATTACTGTGTATCAAGCTAGAACTTTTTTTT	354
Db	246	CATAGTTTAAAGCAATATCTGAGTTCATTACTGTGTATCAAGCTAGAACTTTTTTTT	305
OY	355	CAATCTGGGTTACTCAACCTGTTTATAGACAGATATGGGCACTAGCTGACATGATAT	414
Db	306	CAATCTGGGTTACTCAACCTGTTTATAGACAGATATGGGCACTAGCTGACATGATAT	365
OY	415	GATGGAAGATGAGATCAAGTGGAGTTTCCATAGCTATGAACCTTATCAACTGAAGCTA	474
Db	366	GATGGAAGATGAGATCAAGTGGAGTTTCCATAGCTATGAACCTTATCAACTGAAGCTA	425
OY	475	CAAGATATCAGTACCTCTGCACT	500
Db	426	CAAGATATCAGTACCTCTGCACT	451

RESULT 13
AAZ34573
ID AAZ34573 standard; cDNA; 2079 BP.

DT	01-FEB-2000	(first entry)
XX		
DE	Human SH3D1A	cdna clone 5.

KM	SH3GLA gene; human; Down's syndrome; leukæmia; cancer;
KM	megakaryocytic abnormality; myeloproliferative disorder;
KM	platelet disorder; neural disorder; thrombocytopænia;
KM	haematopoietic disorder; cognitive dysfunction; microcephaly;
KM	lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss
XX	
OS	Homo sapiens.

XX	Key	Location/Qualifiers
XX	Key	136..2079
XX	CDs	/**tag= a
XX	FT	
XX	FT	
XX	PN	MO9953062-A2.
XX	PD	
XX	PD	21-OCT-1999.
XX	PF	
XX	PF	16-APR-1999; 99WO-US008371.
XX	PR	
XX	PR	16-APR-1998; 98US-0082007P.
XX	PA	(CEDA-) CEDARS SINAI HEALTH SYSTEM.
XX	PI	
XX	PI	Korenberg JR, Chen X;
XX	DR	WPI, 1999-633829/54.
XX	DR	P-PSDB; AAY32157.
XX	PT	
XX	PT	Nucleic acid from the human SH3D1A gene and its products, useful for the
XX	XX	diagnosis and treatment of myeloproliferative disorders and leukemia.
XX	PS	Claim 2; Fig 12; 99pp; English.
XX	XX	
XX	XX	This is the nucleotide sequence of a cDNA clone, termed clone 5,
XX	XX	corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
XX	XX	contributes to the development of platelets and the pathogenesis of
XX	XX	leukemias, both in general and in particular those involving the
XX	XX	megakaryocytic lineage. The SH3D1A gene maps to the small candidate
XX	XX	region for low platelets on chromosome 21. Sequencing of 5 different
XX	XX	sizes of cDNA clones from foetal brain (see AA234570-74) suggests that at
XX	XX	least 3 isoforms exist. The invention provides methods for the diagnosis
XX	XX	and treatment of megakaryocytic abnormality, myeloproliferative disorder,
XX	XX	platelet disorder, acute leukemia, neutral disorders, thrombocytopenia,
XX	XX	platelet disorder on chromosome 21, low platelets in deletion for 21,
XX	XX	association of gains in chromosome 21 with leukemias, neutral
XX	XX	abnormalities, dysfunctions and disorders including brain malformations
XX	XX	and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
XX	XX	and colpocephaly. Methods are also provided for: suppressing cells unable
XX	XX	to regulate themselves; screening for a somatic alteration in the SH3D1A
XX	XX	gene; monitoring the progress and adequacy of a treatment; monitoring
XX	XX	tumour risk progress or megakaryocytic abnormality, myeloproliferative
XX	XX	disorder, haematopoietic disorder, platelet disorder or leukemia; and
XX	XX	treatment of a subject (including a prenatal subject) having
XX	XX	megakaryocytic abnormality, myeloproliferative disorder, platelet
XX	XX	disorder, leukemia or neutral disorder using a nucleic acid that
XX	XX	expresses SH3D1A or its antisense nucleic acid
XX	XX	
XX	XX	Sequence 2079 BP; 700 A; 451 C; 495 G; 433 T; 0 U; 0 Other;
XX	XX	
XX	XX	Query Match 85.6%; Score 428; DB 2; Length 2079;
XX	XX	Best Local Similarity 100.0%; Pred. No. 2.3e-120;
XX	XX	Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX	XX	
XX	XX	73 CCGGAGATGATGTCGGGAGCTGCTGCTCTCCCGAGCGCGCTGAGCGCACT 132
XX	XX	1 CCGGAGATGATGTCGGGAGCTGCTGCTCTCCCGAGCGCGCTGAGCGCACT 60
XX	XX	
XX	XX	133 GATTTGTCTCTGGGGGGGAGGCGCGAGCCCGGAGATGAGGCGTTCATTAGCAAGT 192
XX	XX	61 GATTTGTCTCTGGGGGGGAGGCGCGAGCCCGGAGATGAGGCGTTCATTAGCAAGT 120
XX	XX	
XX	XX	133 AAAAGTACAGAAACCTGAGTTCAGACACCTTTGGTGGGAGCGCTGGATATCTGG 252
XX	XX	121 AAAAGTACAGAAACCTGAGTTCAGACACCTTTGGTGGGAGCGCTGGATATCTGG 180
XX	XX	
XX	XX	253 GCCATTAATCTGAGAGAAAGAGCGAGCATGATTCAGACAGTTCATTGTTAAAGCCATA 312
XX	XX	181 GCCATTAATCTGAGAGAAAGAGCGAGCATGATTCAGACAGTTCATTGTTAAAGCCATA 240
XX	XX	
XX	XX	313 TCTGATTCATTAATCTGAGTATTAAGTAAAGCTTTTTCATCTGGGTTACTCTCA 372
XX	XX	

Db	241	TCTGATTCATTAATCTGAGTATTAAGTAAAGCTTTTTCATCTGGGTTACTCTCA 300
Qy	373	CTTTTTCATTAATCTGAGTATTAAGTAAAGCTTTTTCATCTGGGTTACTCTCA 432
Db	301	CTTTTTCATTAATCTGAGTATTAAGTAAAGCTTTTTCATCTGGGTTACTCTCA 360
Qy	433	GTGAGTTCATTAATCTGAGTATTAAGTAAAGCTTTTTCATCTGGGTTACTCTCA 492
Db	361	GTGAGTTCATTAATCTGAGTATTAAGTAAAGCTTTTTCATCTGGGTTACTCTCA 420
Qy	493	TCTGCACT 500
Db	421	TCTGCACT 428
XX	XX	
XX	XX	RESULT 14
XX	XX	AA239008
XX	XX	AA239008 standard; cDNA; 5082 BP.
XX	XX	AA239008;
XX	XX	28-FEB-2000 (first entry)
XX	XX	Mouse Eesl full length cDNA sequence.
XX	XX	
XX	XX	Mouse; murine; Eesl; Ees2; endocytosis; vesicular trafficking;
XX	XX	regulation; actin cytoskeleton; detection; cancer; infection;
XX	XX	EH-domain and SH3-domain regulator of endocytosis; anticancer;
XX	XX	antiproliferative; antiviral; ss.
XX	XX	
XX	XX	Mus sp.
XX	XX	WO9955728-A2.
XX	XX	04-NOV-1999.
XX	XX	27-APR-1999; 99WO-CA000375.
XX	XX	
XX	XX	27-APR-1998; 98CA-02230201.
XX	XX	05-FEB-1999; 99US-0118739P.
XX	XX	(HSCR-) HSC RES & DEV LP.
XX	XX	
XX	XX	Egan SE, Wang W, Sengar A;
XX	XX	WPI, 2000-052892/04.
XX	XX	P-PSDB; AAY57444.
XX	XX	
XX	XX	New nucleic acid encoding Eesl and 2 proteins, involved in regulation of
XX	XX	endocytosis, used e.g. for treating cancer or preventing viral infection.
XX	XX	
XX	XX	Claim 6; Page 38-40; 99pp; English.
XX	XX	
XX	XX	The present sequence encodes mouse Eesl. The present invention
XX	XX	specifically describes mammalian Eesl and 2 proteins (I) and their splice
XX	XX	variants (Bse = EH-domain and SH3-domain regulator of endocytosis). (I)
XX	XX	are involved in regulation of clathrin-mediated endocytosis (as a complex
XX	XX	with Ees1 protein), vesicular trafficking and actin cytoskeleton.
XX	XX	Generally (I) (or its (ant)agonists, mimetics, fragments and inactive
XX	XX	mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
XX	XX	polynucleotide; agents that downregulate expression of Ees genes or
XX	XX	antagonists of an Ees binding partner are used to treat diseases
XX	XX	associated with undesirable endocytosis and resulting changes in cellular
XX	XX	function. Particularly overexpression of Eesl is used to block clathrin-
XX	XX	mediated endocytosis in vivo or in cell cultures, while administration of
XX	XX	(I) is used to promote endocytosis of selected cells. (Ant)agonists of
XX	XX	(I) or Ab are used to suppress abnormal proliferation of cells that can
XX	XX	be stimulated to proliferate by a growth factor receptor; and similar
XX	XX	compounds (also inactive Ees mutants) can be used to prevent viral
XX	XX	infection. Endocytosis may also be regulated, in vivo or in cell
XX	XX	cultures, by forming an Ees-Ees1 complex, then binding dynam to the
XX	XX	complex. Generally conditions that can be treated include cancer;
XX	XX	abnormal cell division or migration; viral infection; or abnormal

receptor signalling, tissue development or synaptic transmission

Sequence 5082 BP, 1363 A; 1253 C; 1391 G; 1075 T; 0 U; 0 Other;
Query Match 71.4%; Score 357; DB 3; Length 5082;
Best Local Similarity 85.5%; Pred. No. 2.1e-98;
Matches 425; Conservative 0; Mismatches 60; Indels 12; Gaps 2;

```
QY 15 GTACGGCGGCTCGCGAGAGAAATCCGAGCGGGCTCCGGAGCGA-----CAG 63
DB 56 GTACGGCGGCTCGCGAGAGAGAAATCCGAGCGGGCTCCGGAGCGAGAGAGAG 115
QY 64 AGAGCGCGCGGGGATGTGTGCGGGGCTGCGGCTCCCTCCGACGGCGCTG 123
DB 116 GCGGCGCGGGGATGTGTGCGGGGCTGCGGCTCCCTCCGACGGCGCTG 174
QY 124 AGCGGCACTGATTTGTCCCTGGGCGGCGGAGCCGCGCGAGATGAGCGTGGAT 183
DB 175 GCGTGCATGATTTGTGTGAGAGGGGCGCGCGGACCCGCGAGATGAGCGTGGAT 234
QY 184 TAGCAAGGTAAAGTACAGAACCATGCTCAGTTTCCAAACCTTTGGTGGAGCTG 243
DB 235 CAGCAAGGTGAACGTAATAGAACCATGCTCAGTTTCCAAACCTTTGGTGGAGCTG 294
QY 244 GATATCTGGGCGCATTAAGTGTAGAGAAAGAGCAAGCATGATCCATAGTTTA 303
DB 295 GATGCTGGGCGCATTAAGTGTAGAGAAAGAGCAAGCATGATCCATAGCTG 354
QY 304 AAGCCAAATATCTGATTCATTAAGTGTATGATCAAGGTAGAACTTTTTCATCTGG 363
DB 355 AAGCCAAATAGCGGATTTTATTAAGTGTATGATCAAGGTAGAACTTTTTCATCTGG 414
QY 364 TTACCTCAACTGTTTATGCAACAGATATGGCACTAGCTGACATGATATGATGAGA 423
DB 415 TTACCTCAACTGTTTATGCAACAGATATGGCACTAGCTGACATGATATGATGAGA 474
QY 424 ATGATCAAGTGTAGTTTCCATGCTATGAACTTATCAACTGAACTCAAGATAT 483
DB 475 ATGATCAAGTGTAGTTTCCATGCTATGAACTTATCAACTGAACTCAAGATAT 534
QY 484 CAGCTACCTCTGCACT 500
DB 535 CAGCTACCTCTCACAACACT 551
```

RESULT 15

AAZ39024
ID AAZ39024 standard; cDNA; 5738 BP.

AC AAZ39024;

DT 28-FEB-2000 (first entry)

DE Mouse Esell cDNA sequence.

XX Mouse; murine; Esel; Ees2; endocytosis; vesicular trafficking;

KW regulation; actin cytoskeleton; detection; cancer; infection;

KW EH-domain and SH3-domain regulator of endocytosis; anticancer;

KW antiproliferative; antiviral; ss.

XX Mus sp.

XX W09955728-A2.

XX 04-NOV-1999.

XX 27-APR-1999; 99WO-CA000375.

XX 27-APR-1998; 98CA-02230201.

XX 05-FEB-1999; 99US-0118739P.

XX (HSCR-) HSC RES & DEV LP.

PI Egan SE, Wang W, Sengar A;
XX WPI; 2000-052802/04.
DR P-PSDB; MAY57449.

PT New nucleic acid encoding Esel and 2 proteins, involved in regulation of
XX endocytosis, used e.g. for treating cancer or preventing viral infection.
PS Claim 6; Page 56-59; 99p; English.

The present invention specifically describes mammalian Esel and 2
proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
regulator of endocytosis). (I) are involved in regulation of clathrin-
mediated endocytosis (as a complex with Espl protein), vesicular
trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
sequences antisense to the (I) polynucleotide; agents that downregulate
expression of Ese genes or antagonists of an Ese binding partner are used
to treat diseases associated with undesirable endocytosis and resulting
changes in cellular function. Particularly overexpression of Esel is used
to block clathrin-mediated endocytosis in vivo or in cell cultures, while
administration of (I) is used to promote endocytosis of selected cells.
(Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
cells that can be stimulated to proliferate by a growth factor receptor;
and similar compounds (also inactive Ese mutants) can be used to prevent
viral infection. Endocytosis may also be regulated, in vivo or in cell
cultures, by forming an Ese-Espl complex, then binding dynamin to the
complex. Generally conditions that can be treated include cancer;
abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission. The
CC present sequence represents mouse Esel cDNA sequence

Sequence 5738 BP, 1556 A; 1456 C; 1613 G; 1113 T; 0 U; 0 Other;

Query Match 71.4%; Score 357; DB 3; Length 5738;
Best Local Similarity 85.5%; Pred. No. 2.2e-98;
Matches 425; Conservative 0; Mismatches 60; Indels 12; Gaps 2;

```
QY 15 GTACGGCGGCTCGCGAGAGAAATCCGAGCGGGCTCCGGAGCGA-----CAG 63
DB 57 GTACGGCGGCTCGCGAGAGAGAAATCCGAGCGGGCTCCGGAGCGAGAGAGAG 116
QY 64 AGAGCGCGCGGGGATGTGTGCGGGGCTGCGGCTCCCTCCGACGGCGGCTG 123
DB 117 GCGGCGCGGGGATGTGTGCGGGGCTGCGGCTCCCTCCGACGGCGGCTG 175
QY 124 AGCGGCACTGATTTGTCCCTGGGCGGCGGAGCCGCGCGAGATGAGCGCTGAT 183
DB 176 GCGTGCATGATTTGTGTGAGAGGGGCGCGCGGACCCGCGAGATGAGCGCTGAT 235
QY 184 TAGCAAGGTAAAGTACAGAACCATGCTCAGTTTCCAAACCTTTGGTGGAGCTG 243
DB 236 CAGCAAGGTGAACGTAATAGAACCATGCTCAGTTTCCAAACCTTTGGTGGAGCTG 295
QY 244 GATATCTGGGCGCATTAAGTGTAGAGAAAGAGCAAGCATGATCCATAGTTTA 303
DB 296 GATGCTGGGCGCATTAAGTGTAGAGAAAGAGCAAGCATGATCCATAGCTG 355
QY 304 AAGCCAAATATCTGATTCATTAAGTGTATGATCAAGGTAGAACTTTTTCATCTGG 363
DB 356 AAGCCAAATAGCGGATTTTATTAAGTGTATGATCAAGGTAGAACTTTTTCATCTGG 415
QY 364 TTACCTCAACTGTTTATGCAACAGATATGGCACTAGCTGACATGATATGATGAGA 423
DB 416 TTACCTCAACTGTTTATGCAACAGATATGGCACTAGCTGACATGATATGATGAGA 475
QY 424 ATGATCAAGTGTAGTTTCCATGCTATGAACTTATCAACTGAACTCAAGATAT 483
DB 476 ATGATCAAGTGTAGTTTCCATGCTATGAACTTATCAACTGAACTCAAGATAT 535
QY 484 CAGCTACCTCTGCACT 500
DB 536 CAGCTACCTCTCACAACACT 552
```

Fri Jul 2 10:23:33 2004

Search completed: July 1, 2004, 11:40:53
Job time : 393 secs

us-09-720-934-1_copy_1_500.rng

```

RESULT 2
US-09-338-933--5
; Sequence 5, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; TITLE OF INVENTION: OVARIAN CANCER
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338,933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ. ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-338-933--5

```

	Query Match	27.6%	Score 137.8;	DB 4;	Length 531;
	Best Local Similarity	64.7%;	Pred. No. 1.1e-33;		
	Matches 205;	Conservative 0;	Mismatches 112;	Indels 0;	Gaps 0;
QY	175	GGCGTCGATTTGCAAGGTAAAGTAAACGAACCAATGGCTCGATTTCCAAACCTTTTGGT	234		
Db	9	GGCTGAGAGCGTCGCAAGAAAGATCAGATTCATGATGGCTCAGTTTCCACAGCATTGAT	68		
QY	235	GGCAGCGCTGGATATCTGGGCCATACCTGTAGAGAAAGCGAAGCATGATGACAGTTG	294		
Db	69	GGAGGGCCAAATATAGTGGCTATTACATCTGAAGACCTACTAAGCAATGATAAACAGTTT	128		
QY	295	CATAGTTTAAAGCAATATCTGATATCTACTGGTGTCAAGCTAGAACTTTT	354		
Db	129	GATTAACCTCAAACTTTCAGAGGTTACATTAAGGTGATCAAGCCCGTACTTTTTCCTA	188		
QY	355	CAATCTGGGTACTCTCAACTCTTTTATGACAGATATGGGCACTAGCTGACATGATAT	414		
Db	189	CAGTCAAGTCTGGCGGGCCCGTTTATGCTGAATATATGGGCTTATCGACTGAAACAG	248		
QY	415	GATGGAAGATGATCAAGTGAAGTTTCCATAGCTATGAATCTTATCAACTGAAGCTA	474		
Db	249	GATGGGAAGATGGACCGACCAAGATCTCTTACTCTATGAATCATCATGATTAAGTTG	308		
QY	475	CAAGSATATGACTACC	491		
Db	309	CAGGGCCAAACAGCTGCC	325		

```

RESULT 3
US-09-215-681-5
/ Sequence 5, Application US/09215681A
/ Patent No. 6528253
/ GENERAL INFORMATION:
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: Fridakis, Tony N.
/ APPLICANT: King, Gordon B.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
/ TITLE OF INVENTION: OF OVARIAN CANCER
/ FILE REFERENCE: 210121.463
/ CURRENT APPLICATION NUMBER: US/09/215,681A
/ CURRENT FILING DATE: 1998-12-17
/ NUMBER OF SEQ. ID NOS: 310
/ SOFTWARE: PasteSeq for Windows Version 3.0
/ SEQ. ID NO 5
/ LENGTH: 531
/ TYPE: DNA
/ ORGANISM: Homo sapien
/ US-09-215-681-5

```

Query Match	27.6%	Score 137.8,	DB 4,	Length 531,
Best Local Similarity	64.7%	Fed. No.1.e-33,		
Matches 205; Conservative	0;	Mismatches 117;	Indels 0;	Gaps 0;

Oy	175	GGCGTCGATTTGCAAGGTAAAAAGTAAACGAACCATGGGCTCAAGTTCCAAACCTTTGGT	234
Db	9	GGCTGAGAGCTGCAAGAAAGAAAGTCAAGATCATGATGGCTCAAGTTCCACAGCGATGAAT	68
Oy	235	GGCAAGCTCGATATCTGGGCCATAACTGTAGAGAAAGAGCAAGCATGATGACAGTTTC	294
Db	69	GGAGGGCCAAATATCTGGGCTATTACATCTGAAGAACCTACACAGCATATATAACAGTTT	128
Oy	295	CATAGTTTAAAGCCAAATCTCGATTCAATTACTGGGATCAAGCTGAAACTTTTCTTTT	354
Db	129	GATAACCTCCAAACCTTCAAGAGGTTACATTAACAGGTGATCAAGCCGCTACTTTTTCCTA	188
Oy	355	CAATCTGGGTTAAGCTCAACCTGTTTTAGACACAGATATGGGCACTAGCTACATGAATTAAT	414
Db	189	CAGTCAGGCTCGCCGGCCCCGGTTTTTAAGCTGAAATATGGGCCCTTTACAAATCTGAAACAAG	248
Oy	415	GATGGAAGATGATGATCAAGTGAAGTTTCCATAGCTATGAACCTATCAACCTGAAGCTA	474
Db	249	GATGGGAAGATGGACACGACGAAGGTTCTCTATAGCTATGAACCTCATCAAGTTTAAAGTTG	308
Oy	475	CAAGATATCAGCTTACC	491
Db	309	CAGGCGCCACAGCTGCC	325

```

RESULT 4
US-09-216-003A-5
Sequence 5, Application US/09216003A
Patent No. 6670463
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Fridakis, Tony N.
APPLICANT: King, Gordon E.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
FILE REFERENCE: 210121.462
CURRENT APPLICATION NUMBER: US/09/216.003A
CURRENT FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 310
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 531
TYPE: DNA
ORGANISM: Homo sapiens
US-09-216-003A-5

```

Query Match	27.6%	Score 137.8	DB 4	Length 531
Best Local Similarity	64.7%	Pred. No. 1.1e-33		
Matches 205; Conservative	0	Mismatches 112	Indels 0	Gaps 0

QY	175	GGGCTGCATTAGCAAGGTAAAGTAACGAGACCATGGCTCACTTTCCAAACCTTTGGT	234
Db	9	GGCTGAAGAGCTCGAAGAAAGAGTCCGAGTACATGATGGCTCAAGTTTCCACAGCGATGAAT	68
QY	235	GGCAGCGCTGGATATCTGGGCCATATCTGTAGAGGAAAGACGAAGCATGATCAGCAGTTT	294
Db	69	GGAGGGCCAAATATGTGGGCTATTACATCTGAAGAACGTACTACAGCATGATTAACAGTTT	128
QY	295	CATAGTTTAAAGCCAAATATCTGGATTCATTTCTGGTGGATCAGGCTAGAAACTTTTTTTT	354
Db	129	GATTAACCTCMAAAGCTTTCAGAGAGTTACATTAACAGGTGATCAAGCCCGTACTTTTTTCCTA	188
QY	355	CAATCTGGGTTAAGCTCAAAGCTGTTTTAGCAACAGATATGGGCACTAGCTGACATGAATAAT	414
Db	189	CAGTCAGGTCCTCCGGGCCCGGTTTAACTGAAATAATGGGCTTATCGAGATCTGAACAAG	248
QY	415	GATGGAAGAAATGATTCAAAGTGGAGTTTCCATAGCTATGAAACTATCAAACTGAAGCTA	474
Db	249	GATGGGAAGATGGACACGACGAAGAGTTCTCTATAGCTATGAAACTCATCAAGTTAAAGTTG	308
QY	475	CAAGGATATCAGCTACC	491
Db	309	CAGGGCCACAGCTGCC	325

```
RESULT 5
US-09-404-879A-72
; Sequence 72, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C3
; CURRENT APPLICATION NUMBER: US/09/404.879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-404-879A-72
```

```
Query Match      27.6%; Score 137.8; DB 4; Length 2017;
Best Local Similarity 64.7%; Pred. No. 2.3e-33;
Matches 205; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 175 GGCCTGATTAGCAAGGTAAAGTAAACCAATGCTCAGTTTCCAAACCTTTTGGT 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9 GGCCTGAGAGCTGCAAGAAAGAGTCAAGATCATGATGCTCAGTTTCCCAAGCCATGAT 68

QY 235 GGCAGCTCGATATCTGGGCCATTAAGTGAAGAAAGCAAGCATGATCAGCATGTC 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 69 GGAGGGCCAAATATGTGGCTATTACATCTGAAGAACTACTAAGCAATAAACAGTTT 128

QY 295 CATAGTTAAAGCCAAATATCTGGATTCTATTCTGGTATCAAGCTAAGAACTTTT 354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 129 GATTAACCTCAACCTTTCAGAGGTTACATTAACAGGTGATCAAGCCGTACTTTTCCCTA 188

QY 355 CAATCTGGGTTAAGCTCAACCTGTTTTCACAGATATGAGCACTGATGATGAATAT 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 189 CAGTAGGCTGCGCCGCCGCCGTTTAACTGAAATATGGCCCTTATCAGATCTGAACAAG 248

QY 415 GATGGAAGATGATCAAGTGAAGTTTCCATAGCTATGAATCAACTTAAGAGCTA 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 249 GATGGGAAGATGACCAAGCAAGTTCTCTATAGCTATGAATCATCAAGTTAAAGTTG 308

QY 475 CAAGGATATCAAGCTACC 491
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 309 CAGGGCCCAACAGCTGCC 325
```

```
RESULT 6
US-09-338-933-72
; Sequence 72, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; TITLE OF INVENTION: OVARIAN CANCER
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338.933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-72
```

Query Match 27.6%; Score 137.8; DB 4; Length 2017;

```
Best Local Similarity 64.7%; Pred. No. 2.3e-33;
Matches 205; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 175 GGCCTGATTAGCAAGGTAAAGTAAACCAATGCTCAGTTTCCAAACCTTTTGGT 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9 GGCCTGAGAGCTGCAAGAAAGAGTCAAGATCATGATGCTCAGTTTCCCAAGCCATGAT 68

QY 235 GGCAGCTCGATATCTGGGCCATTAAGTGAAGAAAGCAAGCATGATCAGCATGTC 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 69 GGAGGGCCAAATATGTGGCTATTACATCTGAAGAACTACTAAGCAATAAACAGTTT 128

QY 295 CATAGTTAAAGCCAAATATCTGGATTCTATTCTGGTATCAAGCTAAGAACTTTT 354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 129 GATTAACCTCAACCTTTCAGAGGTTACATTAACAGGTGATCAAGCCGTACTTTTCCCTA 188

QY 355 CAATCTGGGTTAAGCTCAACCTGTTTTCACAGATATGAGCACTGATGATGAATAT 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 189 CAGTAGGCTGCGCCGCCGCCGTTTAACTGAAATATGGCCCTTATCAGATCTGAACAAG 248

QY 415 GATGGAAGATGATCAAGTGAAGTTTCCATAGCTATGAATCAACTTAAGAGCTA 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 249 GATGGGAAGATGACCAAGCAAGTTCTCTATAGCTATGAATCATCAAGTTAAAGTTG 308

QY 475 CAAGGATATCAAGCTACC 491
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 309 CAGGGCCCAACAGCTGCC 325
```

```
RESULT 7
US-09-215-681-72
; Sequence 72, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Fradakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; TITLE OF INVENTION: OF OVARIAN CANCER
; FILE REFERENCE: 210121.463
; CURRENT APPLICATION NUMBER: US/09/215.681A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-215-681-72
```

```
Query Match      27.6%; Score 137.8; DB 4; Length 2017;
Best Local Similarity 64.7%; Pred. No. 2.3e-33;
Matches 205; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 175 GGCCTGATTAGCAAGGTAAAGTAAACCAATGCTCAGTTTCCAAACCTTTTGGT 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9 GGCCTGAGAGCTGCAAGAAAGAGTCAAGATCATGATGCTCAGTTTCCCAAGCCATGAT 68

QY 235 GGCAGCTCGATATCTGGGCCATTAAGTGAAGAAAGCAAGCATGATCAGCATGTC 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 69 GGAGGGCCAAATATGTGGCTATTACATCTGAAGAACTACTAAGCAATAAACAGTTT 128

QY 295 CATAGTTAAAGCCAAATATCTGGATTCTATTCTGGTATCAAGCTAAGAACTTTT 354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 129 GATTAACCTCAACCTTTCAGAGGTTACATTAACAGGTGATCAAGCCGTACTTTTCCCTA 188

QY 355 CAATCTGGGTTAAGCTCAACCTGTTTTCACAGATATGAGCACTGATGATGAATAT 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 189 CAGTAGGCTGCGCCGCCGCCGTTTAACTGAAATATGGCCCTTATCAGATCTGAACAAG 248

QY 415 GATGGAAGATGATCAAGTGAAGTTTCCATAGCTATGAATCAACTTAAGAGCTA 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 249 GATGGGAAGATGACCAAGCAAGTTCTCTATAGCTATGAATCATCAAGTTAAAGTTG 308
```

QY 475 CAAGATATCAGCTACC 491
DB 309 CAGGCCAACAGCTGCC 325

RESULT 8

US-09-216-003A-72
; Sequence 72, Application US/09216003A
; Patent No. 6670463
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Fridakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
; FILE REFERENCE: 210121.462
; CURRENT APPLICATION NUMBER: US/09/216.003A
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-216-003A-72

Query Match 27.6%; Score 137.8; DB 4; Length 2017;

Best Local Similarity 64.7%; Pred. No. 2.3e-33;
Matches 205; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 175 GCGCTGATTAGCAGTAATAAGTAAACAGAACATGCGCTCAGTTCCAAACCTTTGGT 234
DB 9 GCGCTGAGCTGAGCAAGAAAGTACAGATCATATGAGTCTCAGTTCCACAGCGATGAAT 68
QY 235 GGAGCGCTGGATATCTGGCGCACTAGTAGAGAAAGCGAAGCATATCAGCAGTTTC 294
DB 69 GGAGGCGCAAAATATGCGCTATTAATCATCTGAAAGACGACTAGATGATTAACAGTTT 128
QY 295 CATAGTTTAAAGCAATATCTGATTCATTAATCTGATGATCAAGCTGAAACTTTTTTTT 354
DB 129 GATTAACCTCAACCTTCAGAGAGTTTACATTAACAGGTGATCAAGCCGTTCTTTTCCGA 188
QY 355 CATCTGGGTACTCACTCACTGTTTAAAGACAGATATGCGCACTAGCTGACATGAATAT 414
DB 189 CAGTCAAGGCTCTCGCGCCCGCTTTTAAAGTATGCGCTTATCAATCTGAACAAG 248
QY 415 GATGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 474
DB 249 GATGGGAAGATGACCAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 308
QY 475 CAAGATATCAGCTACC 491
DB 309 CAGGCCAACAGCTGCC 325

RESULT 9

US-09-404-879A-60
; Sequence 60, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404.879A
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapien

US-09-404-879A-60

Query Match 26.7%; Score 133.6; DB 4; Length 480;

Best Local Similarity 66.9%; Pred. No. 2.2e-32;
Matches 190; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 208 ATGGCTCAGTTTCCAAACCTTTTGGTGGCAGCGCTGGATATCTGGCGCATATCTGTAG 267
DB 14 ATGGCTCAGTTTCCCAAGCGATGAATGAGGCGCAAAATATGCGCTATTAATCTGAA 73
QY 268 GAAAGCGAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327
DB 74 GAACGATCAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 133
QY 328 GGTATCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 387
DB 134 GGTATCAAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 193
QY 388 ATATGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
DB 194 ATATGGCGCTTATCAATCTGAAACAGATGAGGCGCAAAATATGCGCTATTAATCTGAA 253
QY 448 GCTATGAAACTTATCAAACTGAAAGCTACAAAGATATCAGCTACC 491
DB 254 GCTATGAAACTCATCAAGTTAAAGTTGACAGGCGCAACAGCTGCC 297

RESULT 10

US-09-338-933-60
; Sequence 60, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338.933
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-60

Query Match 26.7%; Score 133.6; DB 4; Length 480;

Best Local Similarity 66.9%; Pred. No. 2.2e-32;
Matches 190; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 208 ATGGCTCAGTTTCCAAACCTTTTGGTGGCAGCGCTGGATATCTGGCGCATATCTGTAG 267
DB 14 ATGGCTCAGTTTCCCAAGCGATGAATGAGGCGCAAAATATGCGCTATTAATCTGAA 73
QY 268 GAAAGCGAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327
DB 74 GAACGATCAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 133
QY 328 GGTATCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 387
DB 134 GGTATCAAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 193
QY 388 ATATGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
DB 194 ATATGGCGCTTATCAATCTGAAACAGATGAGGCGCAAAATATGCGCTATTAATCTGAA 253
QY 448 GCTATGAAACTTATCAAACTGAAAGCTACAAAGATATCAGCTACC 491
DB 254 GCTATGAAACTCATCAAGTTAAAGTTGACAGGCGCAACAGCTGCC 297

RESULT 11

QY 345 CTTTTTTTCAATCTGGGTTTACCTCAACCTGTTTGAACAGATATGGGCACTAGCTGA 404
DB 1780 GGAGATGCTGGCTCCCAAGTTTCCCAACAGGCTCTGGGCAAGATCTGGAACTAGACCGA 1839
QY 405 CATGAATTAATGATGGAAGAAATGATCAAGTGGAGCTTTTCCATAGCTATGAATCTATCAA 464
DB 1840 CATTGACAGAGATGGCATGTTGGATGACGAGAGTTGGCCCTGGCCCAACCTTATCAA 1899
QY 465 ACTGAAGCTCAAGATATCATGCTACCCCTGCACT 500
DB 1900 AGTCAAGCTAGAGGCGCATGAGCTCCCAAGTGACT 1935

RESULT 14

US-09-312-762A-2
Sequence 2, Application US/09312762A
Patent No. 6552177

GENERAL INFORMATION:

APPLICANT: MIA HOROWITZ ET AL.

TITLE OF INVENTION: EH DOMAIN CONTAINING GENES AND PROTEINS

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Mark M. Friedman c/o Anthony Castorina

STREET: 2001 Jefferson Davis Highway, Suite 207

CITY: Arlington

STATE: Virginia

COUNTRY: United States of America

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk

COMPUTER: Twinhead* Slimote-890TX

OPERATING SYSTEM: MS DOS version 6.2,

SOFTWARE: Word for Windows version 3.11

SOFTWARE: an ASCII file

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/312,762A

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/026,898

FILING DATE: 20 FEB 1998

ATTORNEY/AGENT INFORMATION:

NAME: Friedman, Mark M.

REGISTRATION NUMBER: 33,883

REFERENCE/DOCKET NUMBER: 916/10

TELECOMMUNICATION INFORMATION:

TELEPHONE: 972-3-5625553

TELEFAX: 972-3-5625554

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 3348

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-09-312-762A-2

Query Match

Best Local Similarity 13.8%; Score 69.2; DB 4; Length 3348;

Matches 143; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 235 GGCAGCCTGATATCTGGGCACTAATCTGTAGAGAAAGCGAAGCATGATGCACTTC 294

DB 1379 GGCATTGATGATGTTAGTGGTATGTTGGCAAGACAGCCCACTATATGATCTTC 1438

QY 295 CATAGTTTAAAGCAATATCTGATTCATTACTGTGATCAAGCTTGAATCTTTTTC 354

DB 1439 TACACACTGTCTCTGTCACAGCGCAAGATCAAGGTGCTAATGCCAAGAGAGATGTTG 1498

QY 355 CAATCGGGTTACCTCAACCTTTTGTAGACAGATATGGGCACTAGCTACATGAATAT 414

DB 1499 AAGTCCAGCTGCCCAACACAGTCTGGGGAAGATCTGGAAGTTGGCAGATGGACAG 1558
QY 415 GATGGAAGATGATCAAGTGGAGTTTCCATAGCTATGAATCTATCAAACTGAAGCTA 474
DB 1559 GATGGCTGCTGATGATGACGAGAGTTTGCCTGGCCAAACCTTATCAAGTGAAGCTA 1618
QY 475 CAAGATATGAGTACCTCTGCACT 500
DB 1619 GAGGCCACGAGCTGCCCTGCACT 1644

RESULT 15

US-09-312-762A-3
Sequence 3, Application US/09312762A
Patent No. 6552177

GENERAL INFORMATION:

APPLICANT: MIA HOROWITZ ET AL.

TITLE OF INVENTION: EH DOMAIN CONTAINING GENES AND PROTEINS

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Mark M. Friedman c/o Anthony Castorina

STREET: 2001 Jefferson Davis Highway, Suite 207

CITY: Arlington

STATE: Virginia

COUNTRY: United States of America

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk

COMPUTER: Twinhead* Slimote-890TX

OPERATING SYSTEM: MS DOS version 6.2,

SOFTWARE: Word for Windows version 3.11

SOFTWARE: an ASCII file

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/312,762A

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/026,898

FILING DATE: 20 FEB 1998

ATTORNEY/AGENT INFORMATION:

NAME: Friedman, Mark M.

REGISTRATION NUMBER: 33,883

REFERENCE/DOCKET NUMBER: 916/10

TELECOMMUNICATION INFORMATION:

TELEPHONE: 972-3-5625553

TELEFAX: 972-3-5625554

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 14707

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-09-312-762A-3

Query Match

Best Local Similarity 13.8%; Score 69.2; DB 4; Length 14707;

Matches 143; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 235 GGCAGCCTGATATCTGGGCACTAATCTGTAGAGAAAGCGAAGCATGATGCACTTC 294

DB 9100 GGCATTGATGATGTTAGTGGTATGTTGGCAAGACAGCCCACTATATGATCTTC 9159

QY 295 CATAGTTTAAAGCAATATCTGATTCATTACTGTGATCAAGCTTGAATCTTTTTC 354

DB 9160 TACACACTGTCTCTGTCACAGCGCAAGATCAAGGTGCTAATGCCAAGAGAGATGTTG 9219

QY 355 CAATCGGGTTACCTCAACCTTTTGTAGACAGATATGGGCACTAGCTACATGAATAT 414

DB 9220 AAGTCCAGCTGCCCAACAGTCTGGGGAAGATCTGGAAGTTGGCAGATGGCAAG 9279

QY 415 GATGGAAGATGATCAAGTGGAGTTTCCATAGCTATGAATCTTATCAAACTGAAGCTA 474

Db	9280	GATGACCTGCTGAGATGACGAGAGTTGCCCTGGCCACACCACTTATCAAGTGAAGCTA	9339
Oy	475	CAAGGATATCAGCTACCCCTCTGCACT	500
Db	9340	GAGGGCCACGAGCTGCCCGCTGACT	9365

Search completed: July 1, 2004, 13:02:56
 Job time : 93 secs

This Page Blank (uspto)


```

QY      355  CATCTGGGTTACCTCAACCTGTTTATAGCAAGATTTGGGCACTAGCTGACATGATTAAT 414
      |||
Db      306  CATCTGGGTTACCTCAACCTGTTTATAGCAAGATTTGGGCACTAGCTGACATGATTAAT 365
      |||

QY      415  GATGGAAGATGATCAATCAATGAGTTTCCATAGCTATGAACTTATCAACTGAAGCTA 474
      |||
Db      366  GATGGAAGATGATCAATCAATGAGTTTCCATAGCTATGAACTTATCAACTGAAGCTA 425
      |||

QY      475  CAAGGATATCACTACCTCTGCACT 500
      |||
Db      426  CAAGGATATCACTACCTCTGCACT 451
      |||

RESULT 3
US-09-764-881-55
; Sequence 55, Application US/09764881
; Publication No. US20030125246A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 55
; LENGTH: 568

```

Query Match	Best Local Similarity	Score 345;	DB 10;	Length 568;
Matches 368;	Conservative 0;	Pred.No. 7.2e-102;	Mismatches 3;	Gaps 2;
QY	130	ACTGATTTGTCCTCTGGGGCGGACACCGGACCCCGGAGATGAGCGCTGCATTAGCA	189	
Db	1	ACTGATTTGTCCTCTGGGGCGGCA - CGGGACCCCGCCCGGAATGAGCGCTGCATTAGCA	59	
QY	190	GGTAAAGTACAGAACCATGCTCAAGTTTCCAAACCTTTGGTGGCAGCTGGATATC	249	
Db	60	GGTAAAGTACAGAACCATGCTCAAGTTTCCAAACCTTTGGTGGCAGCTGGATATC	119	
QY	250	TGGGCCATACTGTAGAGGAAAGGCGGAGATATCAGACGTTCATAGTTTAAAGCA	309	
Db	120	TGGGCCATACTGTAGAGGAAAGGCGCA - CATATCAGCAGTTCCATAGTTTAAAGCA	177	
QY	310	ATATCTGATTCATTACTGTGTGATCAAGCTAGAAACTTTTTTTTCATCTGGSTAACT	369	
Db	178	ATATCTGATTCATTACTGTGTGATCAAGCTAGAAACTTTTTTTTCATCTGGSTAACT	237	
QY	370	CAACTGTTTTAGCAGCATGTGGCCACTAGCTGACATGAAATATGATGGAAGATGAT	429	
Db	238	CAACTGTTTTAGCAGCATGTGGCCACTAGCTGACATGAAATATGATGGAAGATGAT	297	
QY	430	CAAGTGAAGTTTCCATAGCTATGAAACTATCAATGAAGCTACAAAGATATCACCTA	489	
Db	298	CAAGTGAAGTTTCCATAGCTATGAAACTATCAATGAAGCTACAAAGATATCACCTA	357	

QY 490 CCCTCTGCACT 500
Db 358 CCCTCTGCACT 368

RESULT 4
US-09-764-875-404
; Sequence 404, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 404
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (536)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (556)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (562)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-875-404

Query Match 69.0%; Score 345; DB 11; Length 568;
Best Local Similarity 99.2%; Pred. No. 7.2e-102;
Matches 368; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

QY 130 ACTGATTTGTCCTGGGGCGGCGAGCGCGACCCCGCGAGATGAGGCGTCGATTAGCAA 189
Db 1 ACTGATTTGTCCTGGGGCGGCGA-CGCGAGCCCGCGAGATGAGGCGTCGATTAGCAA 59
QY 190 GGTAAAGTAACAGAACCATGCTGAGTTCCACACCTTTGGTGGAGCCTGGATATC 249
Db 60 GGTAAAGTAACAGAACCATGCTGAGTTCCACACCTTTGGTGGAGCCTGGATATC 119
QY 250 TGGGCATTAACCTGTAGAGGAAAGAGAGCATGATCAGACGTTCCATGTTAAAGCA 309
Db 120 TGGGCATTAACCTGTAGAGGAAAGAGCA--CATGATCAGACGTTCCATGTTAAAGCA 177
QY 310 ATATCTGATTCATTACTGCTGATCAAGCTAGAACTTTTTTCAATCTGGTTACT 369
Db 178 ATATCTGATTCATTACTGCTGATCAAGCTAGAACTTTTTTCAATCTGGTTACT 237
QY 370 CAACCTGTTTGAACAGATATGGGCACTAGCTGACATGATATGATGAGAAATGAT 429
Db 238 CAACCTGTTTGAACAGATATGGGCACTAGCTGACATGATATGATGAGAAATGAT 297
QY 430 CAAGTGAGTTTCCATAGCTATGAAACTTATCAAACTGAAGCTACAGGATATCAGCTA 489
Db 298 CAAGTGAGTTTCCATAGCTATGAAACTTATCAAACTGAAGCTACAGGATATCAGCTA 357
QY 490 CCCTCTGCACT 500
Db 358 CCCTCTGCACT 368

RESULT 5
US-09-764-881-55

; Sequence 55, Application US/09764881
; Publication No. US20020086821A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (536)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (556)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (562)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-881-55

Query Match 69.0%; Score 345; DB 13; Length 568;
Best Local Similarity 99.2%; Pred. No. 7.2e-102;
Matches 368; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

QY 130 ACTGATTTGTCCTGGGGCGGCGAGCGCGACCCCGCGAGATGAGGCGTCGATTAGCAA 189
Db 1 ACTGATTTGTCCTGGGGCGGCGA-CGCGAGCCCGCGAGATGAGGCGTCGATTAGCAA 59
QY 190 GGTAAAGTAACAGAACCATGCTGAGTTCCACACCTTTGGTGGAGCCTGGATATC 249
Db 60 GGTAAAGTAACAGAACCATGCTGAGTTCCACACCTTTGGTGGAGCCTGGATATC 119
QY 250 TGGGCATTAACCTGTAGAGGAAAGAGAGCATGATCAGACGTTCCATGTTAAAGCA 309
Db 120 TGGGCATTAACCTGTAGAGGAAAGAGCA--CATGATCAGACGTTCCATGTTAAAGCA 177
QY 310 ATATCTGATTCATTACTGCTGATCAAGCTAGAACTTTTTTCAATCTGGTTACT 369
Db 178 ATATCTGATTCATTACTGCTGATCAAGCTAGAACTTTTTTCAATCTGGTTACT 237
QY 370 CAACCTGTTTGAACAGATATGGGCACTAGCTGACATGATATGATGAGAAATGAT 429
Db 238 CAACCTGTTTGAACAGATATGGGCACTAGCTGACATGATATGATGAGAAATGAT 297
QY 430 CAAGTGAGTTTCCATAGCTATGAAACTTATCAAACTGAAGCTACAGGATATCAGCTA 489
Db 298 CAAGTGAGTTTCCATAGCTATGAAACTTATCAAACTGAAGCTACAGGATATCAGCTA 357
QY 490 CCCTCTGCACT 500
Db 358 CCCTCTGCACT 368

RESULT 6
US-10-242-747-55
; Sequence 55, Application US/10242747
; Publication No. US20040005577A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207C1
; CURRENT APPLICATION NUMBER: US/10/242,747
; CURRENT FILING DATE: 2002-09-13

```

1 PRIOR APPLICATION NUMBER: 09/764,1891
2 PRIOR FILING DATE: 2001-01-17
3 PRIOR APPLICATION NUMBER: 60/179,065
4 PRIOR FILING DATE: 2000-01-31
5 PRIOR APPLICATION NUMBER: 60/180,628
6 PRIOR FILING DATE: 2000-02-04
7 PRIOR APPLICATION NUMBER: 60/214,886
8 PRIOR FILING DATE: 2000-06-28
9 PRIOR APPLICATION NUMBER: 60/217,487
10 PRIOR FILING DATE: 2000-07-11
11 PRIOR APPLICATION NUMBER: 60/225,758
12 PRIOR FILING DATE: 2000-08-14
13 PRIOR APPLICATION NUMBER: 60/220,963
14 PRIOR FILING DATE: 2000-07-26
15 PRIOR APPLICATION NUMBER: 60/217,496
16 PRIOR FILING DATE: 2000-07-11
17 PRIOR APPLICATION NUMBER: 60/225,447
18 PRIOR FILING DATE: 2000-08-14
19 PRIOR APPLICATION NUMBER: 60/218,290
20 PRIOR FILING DATE: 2000-07-14
21 Remaining Prior Application data
22 NUMBER OF SEQ ID NOS: 192
23 SOFTWARE: PatentIn Ver. 2.0
24 SEQ ID NO 55
25 LENGTH: 568
26 TYPE: DNA
27 ORGANISM: Homo sapiens
28 FEATURE:
29 NAME/KEY: misc_feature
30 LOCATION: (481)
31 OTHER INFORMATION: n equals a,t,g, or c
32 FEATURE:
33 NAME/KEY: misc_feature
34 LOCATION: (536)
35 OTHER INFORMATION: n equals a,t,g, or c
36 FEATURE:
37 NAME/KEY: misc_feature
38 LOCATION: (556)
39 OTHER INFORMATION: n equals a,t,g, or c
40 FEATURE:
41 NAME/KEY: misc_feature
42 LOCATION: (562)
43 OTHER INFORMATION: n equals a,t,g, or c
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
9
```

Query Match	Similarity	Score	DB	Length
Best Local	Similarity 99.2%	Pfed. No. 7,2e-102		
Matches	368;	Conservative 0;	Mismatches 0;	Indels 3; Gaps 2.
QY	130	ACTGATTTGTCCTCGGCGCGGACAGCGCGAGACCCGCGCGGAGATGAGCGCTGCATTAGCA	189	
Db	1	ACTGATTTGTCCTCGGCGCGCA-CGCGGACCCGCGCGGAGATGAGCGCTGCATTAGCA	59	
QY	190	GGTAAATAACAGAACCATGCTCAGTTTCCAAACACTTTTGGTGGCAGCCTGGATATC	249	
Db	60	GGTAAATAACAGAACCATGCTCAGTTTCCAAACACTTTTGGTGGCAGCCTGGATATC	119	
QY	250	TGGCGCTAATCTGTGAGAGAAAGACGACATGATGCAAGTTCATGATTTAAAGCA	309	
Db	120	TGGCGCTAATCTGTGAGAGAAAGACGACATGATGCAAGTTCATGATTTAAAGCA	177	
QY	310	ATATCTGATTCATTACTGATGATCAAGCTAGAACTTTTTCATCTGGGTACT	369	
Db	178	ATATCTGATTCATTACTGATGATCAAGCTAGAACTTTTTCATCTGGGTACT	237	
QY	370	CAACTGTTTAGCAGAGATATGGGCACTAGCTGACATGAATTAATGATGAGAAATGAT	429	
Db	238	CAACTGTTTAGCAGAGATATGGGCACTAGCTGACATGAATTAATGATGAGAAATGAT	297	
QY	430	CAAGTGGAGTTTCCATAGCTATGAACCTATCAACCTGAAGCTACAAAGATATCAGCTA	489	
Db	298	CAAGTGGAGTTTCCATAGCTATGAACCTATCAACCTGAAGCTACAAAGATATCAGCTA	357	

QY	490	CCCTCTGCACT	500
Db	358	CCCTCTGCACT	368

```

RESULT 7
US-10-158-057-127
; Sequence 127, Application US/10158057
; Publication No. US20040014039a1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1205C1
; CURRENT APPLICATION NUMBER: US/10/158,057
; PRIORITY FILING DATE: 2002-06-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 127
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (536)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (556)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (562)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-158-057-127

```

Query Match	Similarity	99.0%	Score 345	DB 163	Length 568
Best Local	Similarity	99.2%	Pred. No. 7,2e-102		
Matches 368	Conservative	0	Mismatches	0	Indels
					Gaps
					2
QY	130	ACTGATTTGTCCTGGGGGCGGACGGCCGAGCCCGCCGAGATGACGGCTGCATTAGCAA	189		
Db	1	ACTGATTTGTCCTGGGGGCGGCA--CGGGACCCCGCCGAGATGAGGGCTGCATTAGCAA	59		
QY	190	GGTAAAGTAAACAGACCATGGCTCAGTTTCCAAACCTTTTGGTGGCAGCCTGGATATC	249		
Db	60	GGTAAAGTAAACAAACCATGGCTCAGTTTCCAAACCTTTTGGTGGCAGCCTGGATATC	119		
QY	250	TGGGCGCATTAAGTGTAGAGAAAGACGAGCATGATGACAGTTCGCATAGTTTAAAGCCA	309		
Db	120	TGGGCGCATTAAGTGTAGAGAAAGAGCCCA--CATGATGACAGTTCGCATAGTTTAAAGCCA	177		
QY	310	ATATCTGATTCATTACTGCTGTGATCAGACTAGAAACCTTTTTCATCTGGGGTATCCT	369		
Db	178	ATATCTGATTCATTACTGCTGTGATCAGACTAGAAACCTTTTTCATCTGGGGTATCCT	237		
QY	370	CAACCTGTTTATGACACAGATATGGGCGACTACTGACATGATTAATGAGAAATGAT	429		
Db	238	CAACCTGTTTATGACACAGATATGGGCGACTACTGACATGATTAATGAGAAATGAT	297		
QY	430	CAAGTGAAGTTTCCATAGCTATGATAAACTTATCAAAGTGAAGCTACAAAGATATCAGCTA	489		
Db	298	CAAGTGAAGTTTCCATAGCTATGATAAACTTATCAAAGTGAAGCTACAAAGATATCAGCTA	357		
QY	490	CCCTCTGCACT	500		
Db	358	CCCTCTGCACT	368		

```

RESULT 8
US-09-764-881-50
; Sequence 50, Application US/09764881
; Publication No. US20030125246A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PAM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 967
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (659)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (697)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (715)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (869)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (879)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-881-50

Query Match      30.6%; Score 152.8; DB 10; Length 967;
Best Local Similarity 67.5%; Pred. No. 7,8e-39;
Matches 208; Conservative 4; Mismatches 96; Indels 0; Gaps 0

Db
184 TAGCAAGGTAAAGTACAGAACCATGGCTCAGITTCACACACTTTTGGTGCAGCCTG 243
218 TAGCAAGAAAGAACTCAGAGCAATGATGGCTCAGTTTCCACAGCTATGATGAGGGCCA 277
244 GATATCTGGGCCATCTACTGTGAGGAAAGCGAAGCATGATCAGCAGTTCATGTTTA 303
278 AACATGTGKGCATATTAACCTCTGMAAGAGTACTTAAGCATGACAGGCAGTTGTATACCTC 337
304 AAGCCAATATCTGAGTTCATTACTGTGTATCAAGCTAGAACTTTTTCATCTGGG 363
338 AAACCTTAGAGAGGTATCAATAACAGGTATCAAGCACGTAATTTTTCCTACATAGGT 397
364 TTACCTCAACCGTTTTCAGACAGATATGGGCACTAGCTGCATGAAATATGATGAGAGA 423
398 CTCGGCGGCCCTGTTTTCAGTAATATGGGCTTATACAGACTTAACMGKATGGGAAG 457
424 ATGATCAAGTGAAGTTTTCATAGCTATGAATCTTATCAAACTGAAGCTACAGGATAT 483
458 ATGATTCAGCAAGAGTTCCTCATAGCTATGAATCAATCAACTGAAGCTTCAGGCCAA 517
484 CAGCTAAC 491
518 CAGTTGCC 525

RESULT 9
US-09-764-875-411
; Sequence 411, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT202
; CURRENT APPLICATION NUMBER: US/09/764,875

```

```

CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1249
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 411
LENGTH: 967
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (659)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (697)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (715)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (879)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (879)
OTHER INFORMATION: n equals a,t,g, or c
US-09-764-875-411

Query Match      30.6%; Score 152.8; DB 11; Length 967;
Best Local Similarity 67.5%; Fred.No. 7,8e-39;
Matches 208; Conservative 4; Mismatches 96; Indels 0; Gaps 0

QY 184 TAGCAAGTAATAAAGTACAGAACAATGGCTGAGTTTCACAACACTTTTGTTGGCAGCCGTG 243
DB 218 TAGCAAGGAATACTCAGACCATGATGGCTCAGTTTCCACAGCTATGATGAGGCCA 277
QY 244 GATATCTGGCCATPACTACTGTAGAGAAAGAGCAATGATCAGCAGTTCATAGTTTA 303
DB 278 AACATGTGKGCATTACCTCGAAGAAAGTACTAAGCATGACAGGAGTTGATTAACCTC 337
QY 304 AAGCCAATATCTGATTATTAATCTGTGATCAAGCTAGAACCTTTTTTTTCAATCTGG 363
DB 338 AAACCTTAGAGAGTTACATTAACAAGATGATCAAGCACGTAATTTTCTTACATCAGGT 397
QY 364 TTACCTCAACCGTTTGTAGCAAGATATGGGCACTAGCTGCAATGAATATGATGAAGA 423
DB 398 CTGCGCGCCCCGTTTATGCTGAATATGAGGCTTTATCTAGACCTTAACMGATGGGAAG 457
QY 424 ATGATCAAGTGAAGTTTTCATAGCTATGAACCTTATCAAACTGAAGCTACAGGATAT 483
DB 458 ATGATTCAGCAAGAGTTCTCCATACCTATGAAGACTCATAACTGAAGCTTCAAGGCCA 517
QY 484 CAGCTACC 491
DB 518 CAGTTGCC 525

RESULT 10
US-09-764-881-50
Sequence 50, Application US/09764881
Publication No. US20020086821A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic acids, proteins, and Antibodies
FILE REFERENCE: PT207
CURRENT APPLICATION NUMBER: US/09/764,881
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 192
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 50
LENGTH: 967
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
```

NUMBER OF SEQ ID NOS: 384


```
SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 967
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (659)
;   OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (697)
;   OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (715)
;   OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (869)
;   OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (879)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-158-057-40
```

```
Query Match      30.6%; Score 152.8; DB 16; Length 967;
Best Local Similarity 67.5%; Pred. No. 7.8e-39;
Matches 208; Conservative 4; Mismatches 96; Indels 0; Gaps 0;
```

```
QY 184 TAGCAAGTAAAGTAACAGAACCATGGCTCAGTTTCCAAACCTTTGGTGGCAGCCTG 243
DB 218 TAGCAAGTAAAGTAACAGAACCATGGCTCAGTTTCCAAACCTTTGGTGGCAGCCTG 243
QY 244 GATATCTGGGCGCCATTAATCTGATGAGGAAAGAGGAGGAGGAGGAGGAGGAGGAGG 303
DB 278 AACATGCGGTATTAATCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 303
QY 304 AAGCCAAATCTGATTAATCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 363
DB 338 AACCTTCAGAGGATTAATCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 397
QY 364 TTACTCAACCTGTTTGAACAGATATGGGCACTAGCTGACATGATATGATGAGGAGG 423
DB 398 CTGCGCGGCCCTGTTTGAACAGATATGGGCACTAGCTGACATGATATGATGAGGAGG 457
QY 424 ATGATCAAGTGAAGTTTCCATAGCTATGAACTTATCAACTGAACTCAAGGATAT 483
DB 458 ATGATCAAGTGAAGTTTCCATAGCTATGAACTTATCAACTGAACTCAAGGATAT 517
QY 484 CAGCTACC 491
DB 518 CAGTTGCC 525
```

```
RESULT 13
US-10-398-885A-15
; Sequence 15, Application US/10398885A
; Publication No. US20040053282A1
; GENERAL INFORMATION:
;   APPLICANT: Sugita, Yuji
;   APPLICANT: Hashida, Ryoichi
;   APPLICANT: Ogawa, Kaoru
;   APPLICANT: Nagasu, Takeshi
;   APPLICANT: Obayashi, Masaya
;   APPLICANT: Saito, Hirohisa
;   APPLICANT: Takahashi, Eiki
; TITLE OF INVENTION: Method of Testing For Allergic Diseases
; FILE REFERENCE: SHIMIZU-07907
; CURRENT APPLICATION NUMBER: US/10/398,885A
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: PCT/JP01/08937
```

```
PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: JP 2000-314093
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 5828
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
;   OTHER INFORMATION: Synthetic
;   NAME/KEY: CDS
;   LOCATION: (7)..(5052)
;   OTHER INFORMATION:
US-10-398-885A-15
```

```
Query Match      30.6%; Score 152.8; DB 13; Length 5828;
Best Local Similarity 68.5%; Pred. No. 2.5e-38;
Matches 211; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
```

```
QY 184 TAGCAAGTAAAGTAACAGAACCATGGCTCAGTTTCCAAACCTTTGGTGGCAGCCTG 243
DB 19 TATCAAGTAAAGTAACAGAACCATGGCTCAGTTTCCAAACCTTTGGTGGCAGCCTG 78
QY 244 GATATCTGGGCGCCATTAATCTGATGAGGAAAGAGGAGGAGGAGGAGGAGGAGGAGG 303
DB 278 AACATGCGGTATTAATCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 303
QY 304 AAGCCAAATCTGATTAATCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 363
DB 338 AACCTTCAGAGGATTAATCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 397
QY 364 TTACTCAACCTGTTTGAACAGATATGGGCACTAGCTGACATGATATGATGAGGAGG 423
DB 398 CTGCGCGGCCCTGTTTGAACAGATATGGGCACTAGCTGACATGATATGATGAGGAGG 457
QY 424 ATGATCAAGTGAAGTTTCCATAGCTATGAACTTATCAACTGAACTCAAGGATAT 483
DB 458 ATGATCAAGTGAAGTTTCCATAGCTATGAACTTATCAACTGAACTCAAGGATAT 517
QY 484 CAGCTACC 491
DB 518 CAGTTGCC 525
```

```
RESULT 14
US-09-884-441-5
; Sequence 5, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
;   APPLICANT: Algate, Paul A.
;   APPLICANT: Carter, Patrick
;   APPLICANT: Carter, Patrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;   TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C7
; CURRENT APPLICATION NUMBER: US/09/884,441
; CURRENT FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-884-441-5
```

```
Query Match      27.6%; Score 137.8; DB 9; Length 531;
Best Local Similarity 64.7%; Pred. No. 4.3e-34;
Matches 205; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
```

```
QY 175 GGGCTGATTAGCAAGTAAAGTAACAGAACCATGGCTCAGTTTCCAAACCTTTGGT 234
DB 9 GGGCTGAGAGCTGCAAGTAAAGTCAAGATCATGATGAGTTCAGTTTCCCAAGCATGAT 68
```

QY 235 GCGAGCTGTGATATCTGGCCATACTGTAGAGAAAGACGAGCATGACAGCTTC 294
 Db 69 GGAGGCCCAAAATATGTGGGCTATTACATCTGAAGAACGACTAGATGATTAACAGTTT 128
 QY 295 CATAGTTTAAAGCCAAATATCTGCATTCATTACTGTGATCAAGCTGAACCTTTT 354
 Db 129 GATAACCTCAAACTTCAGGAGGTACATACAGGTGATCAAGCCGCTACTTTTCCCTA 188
 QY 355 CATCTGGGTTACCTCAACCTTTTATAGACAGATATGGGCACTAGCTACATGAATAAT 414
 Db 189 CAGTACGGTCTGCGGCGCCGCTTTTATGCTGAATAATG3CCCTTATCAGATCTGAACAAAG 248
 QY 415 GATGGAAGATGAGATCAAGTGTGATTTTCCATAGCTATGAACCTTCAACCTGAAGCTA 474
 Db 249 GATGGGAAGATGAGACGAGCAAGATTTCTATATGCTATGAACCTCATCAAGTTAAAGTTG 308
 QY 475 CAAGATATCAGCTACC 491
 Db 309 CAGGCCCAACAGCTGCC 325

RESULT 15

US-09-907-969-5
 ; Sequence 5, Application US/09907969
 ; Publication No. US20030091580A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Fling, Steven P.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Panger, Gary Richard
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Vedrick, Thomas S.
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Hill, Paul
 ; APPLICANT: Albone, Earl
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE REFERENCE: 210121.462C8
 ; CURRENT APPLICATION NUMBER: US/09/907,969
 ; NUMBER OF SEQ ID NOS: 596
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 531
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-907-969-5

Query Match 27.6%; Score 137.8; DB 10; Length 531;
 Best Local Similarity 64.7%; Pmed.No. 4.3e-34;
 Matches 205; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 175 GCGGTGATTAGCAGGTAAAGTAACAGAACATGCTCAGTTTCCAAACCTTTGGT 234
 Db 9 GGCTGAGAGCTGCAAGAGAGAGTCAAGATCATGATGGCTCACTTCCACAGCGATGAAT 68
 QY 235 GCGAGCTGTGATATCTGGGCTTAACCTGTAGAGAAAGCAGAGCATGATCAGCAAGTTT 294
 Db 69 GGAGGCCCAAAATATGTGGGCTATTACATCTGAAAGACGACTAAGCATGATTAACAGTTT 128
 QY 295 CATAGTTTAAAGCCAAATATCTGCATTCATTACTGTGATCAAGCTAGAAACCTTTT 354
 Db 129 GATAACCTCAAACTTCAGGAGGTACATTAACAGGTGATCAAGCCGCTACTTTTCCCTA 188
 QY 355 CAATCTGGGTTACTCAACCTTTTATAGACAGATATGGGCACTAGCTGACATGAATAAT 414
 Db 189 CAGTACGGTCTGCGGCGCCGCTTTTATGCTGAATAATG3CCCTTATCAGATCTGAACAAAG 248
 QY 415 GATGGAAGATGAGTCAAGTGTGATTTTCCATAGCTATGAACCTTATCAAACTGAAGCTA 474

Db 249 GATGGGAAGATGAGACGAGCAAGATTTCTATATGCTATGAACCTCATCAAGTTAAAGTTG 308
 QY 475 CAAGATATCAGCTACC 491
 Db 309 CAGGCCCAACAGCTGCC 325

Search completed: July 1, 2004, 13:09:09
 Job time : 365 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 12:16:50 ; Search time 2007.31 Seconds
(without alignments)
1900.151 Million cell updates/sec

Title: US-09-720-934-2_COPY_15_102
Perfect score: 459
Sequence: 1 WAITVEERAKHDQGFHSLKP.....KLIKKLQGYQLPSALPPVM 88

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 segs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seg length: 0
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xid
-O=/cgm2.1/USPTO.spool_p/US9720934/runat.30062004.064539.13430/app.query.fasta_1.1386
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0 -LOOPEXT=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US09720934 @CGN 1.1 7509 @runat.30062004.064539.13430 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOUTRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ha:*
2: gb_hhg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_hhg_hum:*
31: em_hhg_inv:*
32: em_hhg_other:*
33: em_hhg_mus:*
34: em_hhg_pin:*
35: em_hhg_rod:*
36: em_hhg_mam:*
37: em_hhg_vrt:*
38: em_sy:*
39: em_hhgo_hum:*
40: em_hhgo_mus:*
41: em_hhgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB	ID	Description
1	459	100.0	877	6	BD124639	BD124639 Primer fo
2	459	100.0	877	6	BD126208	BD126208 Primer fo
3	459	100.0	1299	9	BC039036	BC039036 Homo sapi
4	459	100.0	1676	6	BD127640	BD127640 Primer fo
5	459	100.0	1676	9	AK075290	AK075290 Homo sapi
6	459	100.0	2079	6	BD205036	BD205036 Isolated
7	459	100.0	2131	6	AX880753	AX880753 Sequence
8	459	100.0	2131	6	BD127168	BD127168 Primer fo
9	459	100.0	2131	9	AK027846	AK027846 Homo sapi
10	459	100.0	2131	9	AK074554	AK074554 Homo sapi
11	459	100.0	2131	9	BC058925	BC058925 Homo sapi
12	459	100.0	2199	6	BD205035	BD205035 Isolated
13	459	100.0	5195	6	BD205033	BD205033 Isolated
14	459	100.0	5199	6	BD205033	BD205033 Isolated
15	459	100.0	5287	9	AF064243	AF064243 Homo sapi
16	459	100.0	5361	9	AF114488	AF114488 Homo sapi
17	459	100.0	5458	6	BD205034	BD205034 Isolated
18	459	100.0	6439	9	AF114487	AF114487 Homo sapi
19	459	100.0	7247	9	AF064244	AF064244 Mus muscu
20	441	96.1	799	10	AF525079	AF525079 Mus muscu
21	441	96.1	3723	10	AF132478	AF132478 Mus muscu
22	441	96.1	5145	10	AF132481	AF132481 Rattus no
23	440	95.9	3812	10	AF132672	AF132672 Rattus no
24	440	95.9	4025	10	AF127798	AF127798 Rattus no
25	426	92.8	676	6	AX870076	AX870076 Sequence
26	426	92.8	676	6	BD150138	BD150138 Primer fo
27	409	89.1	4103	5	AF032118	AF032118 Xenopus l
28	389	84.7	481	10	AF468654	AF468654 Mus muscu
29	377	82.1	3390	9	BSM809003	BSM809003 Homo sapi
30	365	79.5	1749	9	BC020921	BC020921 Homo sapi
31	365	79.5	2085	9	BC038963	BC038963 Homo sapi
32	365	79.5	4557	9	AF182199	AF182199 Homo sapi
33	365	79.5	5828	6	BD167848	BD167848 Method fo
34	365	79.5	5828	6	AF248540	AF248540 Homo sapi
35	365	79.5	5938	9	AB033082	AB033082 Homo sapi
36	365	79.5	6091	9	AF182198	AF182198 Homo sapi
37	359	78.2	480	6	BD264675	BD264675 Compositi
38	359	78.2	480	6	AR238078	AR238078 Sequence
39	359	78.2	480	6	AR257619	AR257619 Sequence
40	359	78.2	480	6	AR283665	AR283665 Sequence
41	359	78.2	480	6	AX366293	AX366293 Sequence
42	359	78.2	531	6	BD264620	BD264620 Compositi
43	359	78.2	531	6	AR238023	AR238023 Sequence
44	359	78.2	531	6	AR257564	AR257564 Sequence
45	359	78.2	531	6	AR283610	AR283610 Sequence

RESULT 1

ALIGNMENTS

BD124639
 LOCUS BD124639 877 bp DNA linear PAT 18-SEP-2002
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION BD124639
 VERSION BD124639.1 GI:23219584
 KEYWORDS JP 2002017375-A/70.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y., Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.
 REFERENCE 1 (bases 1 to 877)
 AUTHORS Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y., Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.
 TITLE Primer for synthesizing full-length cDNA and use thereof
 JOURNAL Patent: JP 2002017375-A/70 22-JAN-2002;
 COMMENT HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002017375-A/70
 PD 22-JAN-2002
 PF 07-JUL-2000 JP 2000253172
 PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO ISHII,
 PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI SHINICHI KOJIMA,
 PI TETSUJI OTSUKI, HISASHI KOGA
 PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N5/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key
 FEATURES
 source Location/Qualifiers
 1..877 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,72e-53 Length: 877
 Score: 459.00 Matches: 88
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 DB: 6
 US-09-720-934-2_COPY_15_102 (1-88) x BD124639 (1-877)
 QY 1 TTPAATLTHPRVAGLUGUAGUAGALALYSHISAAPGNGINPHEHISERLEULYSPRO 20
 Db 247 TGGGCGCATTAAGTGAAGAAAGCGAAGCATATGACGATTCATATGTTTAAAGCCA 306
 QY 21 ILESERGLYPHELIETHNGLYASPGINALAARGANPHEPHEGINSERGLYLEUPRO 40
 Db 307 ATATCTGGATTCATTAATCGATCAAGCTAAGAACTTTTTCATCTGGCTTACT 366
 QY 41 GINPROVALLEUALAGNILETPALALEUALAASPWETAENANASPGIYARGMETASP 60
 Db 367 CAACTCTGTTTAAAGACAGATATGGGCACTAGCTACATGAATATGATGGAAGATGAT 426
 QY 61 GINVALGUPHESERILEALAMELYSEULLEYSEULYSEULNGIYTYRGINLEU 80
 Db 427 CAACTGGAGTTTCCATAGCTATGAACCTTATCAAACTGAAGCTACAAAGATATCAGCTA 486
 QY 81 PROSERIALAENUPROVALMET 88
 Db 487 CCCTCTGCACTTCCCTCTGCATG 510
 RESULT 2
 BD126208

BD126208
 LOCUS BD126208 877 bp DNA linear PAT 18-SEP-2002
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION BD126208
 VERSION BD126208.1 GI:23221153
 KEYWORDS JP 2002017375-A/1639.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y., Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.
 REFERENCE 1 (bases 1 to 877)
 AUTHORS Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y., Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.
 TITLE Primer for synthesizing full-length cDNA and use thereof
 JOURNAL Patent: JP 2002017375-A 1639 22-JAN-2002;
 COMMENT HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002017375-A/1639
 PD 22-JAN-2002
 PF 07-JUL-2000 JP 2000253172
 PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO ISHII,
 PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI SHINICHI KOJIMA,
 PI TETSUJI OTSUKI, HISASHI KOGA
 PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N5/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key
 FEATURES
 source Location/Qualifiers
 1..877 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,72e-53 Length: 877
 Score: 459.00 Matches: 88
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 DB: 6
 US-09-720-934-2_COPY_15_102 (1-88) x BD126208 (1-877)
 QY 1 TTPAATLTHPRVAGLUGUAGUAGALALYSHISAAPGNGINPHEHISERLEULYSPRO 20
 Db 247 TGGGCGCATTAAGTGAAGAAAGCGAAGCATATGACGATTCATATGTTTAAAGCCA 306
 QY 21 ILESERGLYPHELIETHNGLYASPGINALAARGANPHEPHEGINSERGLYLEUPRO 40
 Db 307 ATATCTGGATTCATTAATCGATCAAGCTAAGAACTTTTTCATCTGGCTTACT 366
 QY 41 GINPROVALLEUALAGNILETPALALEUALAASPWETAENANASPGIYARGMETASP 60
 Db 367 CAACTCTGTTTAAAGACAGATATGGGCACTAGCTACATGAATATGATGGAAGATGAT 426
 QY 61 GINVALGUPHESERILEALAMELYSEULLEYSEULYSEULNGIYTYRGINLEU 80
 Db 427 CAACTGGAGTTTCCATAGCTATGAACCTTATCAAACTGAAGCTACAAAGATATCAGCTA 486
 QY 81 PROSERIALAENUPROVALMET 88
 Db 487 CCCTCTGCACTTCCCTCTGCATG 510
 RESULT 3
 BC039036
 LOCUS BC039036 1299 bp mRNA linear PRI 24-DEC-2002

US-09-720-934-2_COPY_15_102 (1-88) x BC039036 (1-1259)	
QY	1 TYPAla1leThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
Db	306 TGGGCCATACTCTAGAGGAAAGCGCAGCATGATCAGCAGTTCACATGTTTAAAGCA 365
QY	21 H1SerGlyPheIleThrGlyAspGlnAlaArgAsnGlnPhePheGlnSerGlyLeuPro 40
Db	366 ATATCTGGATTCATTACTGGGTGATCAAGTCAGAACTTTTTCATCTGGGTACT 425
QY	41 GlnProValIleuAlaGlnIleTyrAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
Db	426 CAACCTGTTTATGACACAGATATGGGCACTAGCTGACATGATATGATGAAAGATGAT 485
QY	61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
Db	486 CAACTGGAGTTTCCATAGCTATATAAACTTATCAAACTGAAGCTACAGAGATATCAGCTA 545
QY	81 ProSerAlaLeuProProValMet 88
Db	546 CCTCTGCACCTTCCCTCTGTCATG 569
RESULT 4	
LOCUS	BD127640 1676 bp DNA linear PAT 18-SEP-2002
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.
ACCESSION	BD127640
VERSION	BD127640.1 GI:23222585
KEYWORDS	JP 2002017375-A/3071.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eumetazoa; Primates; Catarrhini; Homiidae; Homo.
AUTHORS	1 (bases 1 to 1676) Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y., Makamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.
TITLE	Primer for synthesizing full-length cDNA and use thereof
JOURNAL	Patent: JP 2002017375-A 3071 22-JAN-2002;
COMMENT	HELIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002017375-A/3071 PD 22-JAN-2002 PF 07-JUL-2000 JP 2000253172 PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO PI ISHII, PI YURI KAWAI,AI MAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI SHINICHI KOJIMA, PI TETSUO OTSUKI,HISASHI KOGA PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC 10' C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key Location/Qualifiers (264) . (1676) . FT CDS Location/Qualifiers 1..1676 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"
ORIGIN	
Alignment Scores:	
Pred. No.:	5,79e-53 Length: 1676
Score:	459.00 Matches: 88
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	Gaps: 0
US-09-720-934-2_COPY_15_102 (1-88) x BD127640 (1-1676)	

QY 1 TTPAaIIEThVaIGluGluarGaIaYSHsASpGInGInPheHisserLeuYsPro 20
 Db 306 TGGGCGCATACTGTAGAGAAAGAGCGAAGCATGATCAGACAGTTTCAATGTTAAAGCCA 365
 QY 21 TIESerGlyPheIIEhrcGlyAspGlnAaIaAgAsnPhePheGInserGlyLeuPro 40
 Db 366 ATATCTGAGTTTCAATTACTGTGTATCAAGCTAGAACCTTTTTTTTTCATCAATCGGGTTACCT 425
 QY 41 GInProValIeuaAGInIIEThPaIaLeuaAaAspMetAsnAsnAspGlyArGMeTAsp 60
 Db 426 CAACCGTTTTCAGCAGATATGGGACACTAGCTGACATGATATGATGAGAAATGGAT 485
 QY 61 GInValGluPheSerIIEaIaMetIySLeuIIEySLeuYsLeuGInGInGlyTyrgInLeu 80
 Db 486 CAAGTGGAGTTTTCATGACTATGAACCTTAACTGAAGCTACAGGATATACGCTA 545
 QY 81 ProSerAlaLeuProProValMet 88
 Db 546 CCTCTGACATTCCTCCCTGTCAATG 569
 RESULT 5
 AK075290
 LOCUS 1676 bp mRNA linear PRI 03-SEP-2002
 DEFINITION Homo sapiens CDNA FLJ90409 fis, clone Y79AA1000778, highly similar to Homo sapiens Intesedctn long form mRNA.
 ACCESSION AK075290
 VERSION AK075290.1 GI:22761283
 KEYWORDS oligo cloning; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1
 AUTHORS Iisoga,T., Oca,T., Nishikawa,T., Hayashi,K., Otsuki,T., Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S., Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Kojima,S., Nagahara,K., Masuno,Y., Ono,T., Okano,K., Yoshikawa,Y., Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and Niinomiya,K.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1676)
 AUTHORS Iisoga,T. and Otsuki,T.
 TITLE Direct Submission
 JOURNAL Submitted (25-MAR-2002) Takao Iisoga, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomic@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo; Laboratory of Genome Structure, Human Genome Center; cDNA 5' - 3' end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).
 FEATURES
 source
 1..1676
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="Y79AA1000778"
 /cell_line="Y79"
 /cell_type="retinoblastoma"
 /clone_id="Y79AA1"
 /note="Cloning vector: pME18SFL3"

ORIGIN

Alignment Scores:
 Pred. No.: 5.79e-53
 Score: 459.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 DB: 9
 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x AK075290 (1-1676)
 QY 1 TTPAaIIEThVaIGluGluarGaIaYSHsASpGInGInPheHisserLeuYsPro 20
 Db 306 TGGGCGCATACTGTAGAGAAAGAGCGAAGCATGATCAGACAGTTTCAATGTTAAAGCCA 365
 QY 21 TIESerGlyPheIIEhrcGlyAspGlnAaIaAgAsnPhePheGInserGlyLeuPro 40
 Db 366 ATATCTGAGTTTCAATTACTGTGTATCAAGCTAGAACCTTTTTTTTTCATCAATCGGGTTACCT 425
 QY 41 GInProValIeuaAGInIIEThPaIaLeuaAaAspMetAsnAsnAspGlyArGMeTAsp 60
 Db 426 CAACCGTTTTCAGCAGATATGGGACACTAGCTGACATGATATGATGAGAAATGGAT 485
 QY 61 GInValGluPheSerIIEaIaMetIySLeuIIEySLeuYsLeuGInGInGlyTyrgInLeu 80
 Db 486 CAAGTGGAGTTTTCATGACTATGAACCTTAACTGAAGCTACAGGATATACGCTA 545
 QY 81 ProSerAlaLeuProProValMet 88
 Db 546 CCTCTGACATTCCTCCCTGTCAATG 569
 RESULT 6
 BD205036
 LOCUS 2079 bp DNA linear PAT 17-JUL-2003
 DEFINITION Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.
 ACCESSION BD205036
 VERSION BD205036.1 GI:33014806
 KEYWORDS JP 2002511267-A/4.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 2079)
 AUTHORS Korenberg,J.R. and Chen,X.N.
 TITLE Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof
 JOURNAL Patent: JP 2002511267-A 4 16-APR-2002;
 CEDARS SINAI HEALTH SYSTEM ET AL
 OS Homo sapiens (human)
 PN JP 2002511267-A/4
 PD 16-APR-2002
 PF 16-APR-1999 JP 2000543610
 PR 16-APR-1998 US 60/082007
 PT JULIE R KORENBERG,XIAO NING CHEN
 PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,
 PC C12Q1/68,
 PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
 CC Isolated SH3 gene relating to myeloproliferative disorders and

leukemia
 CC and utilization thereof.
 FH Key Location/Qualifiers
 FT source 1..2079
 /organism="Homo sapiens (human)"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 7.45e-53
 Score: 459.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 DB: 6
 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BD205036 (1-2079)

QY 1 TTPAlaIleThrValGluGluArgAlaYshHsAspGlnGlnPheHisSerLeuIysPro 20
 DB 178 TGGGCATTAAGTCTAGAGGAAAGAGGAGCATGATCAGAGTTCCATAGTTAAAGCCA 237
 QY 21 IIsErglyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
 DB 238 ATATCTGGATTCTATCTGTGATCAAGCTAGAAACCTTTTTCATCTGGATTACCT 297
 QY 41 GlnProValLeuAlaGlnIleThrPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
 DB 298 CAACCTGTTTACCAAGATATGGGCACTAGCTGACATGATATATGATGGAAGAATGGAT 357
 QY 61 GlnValGluPheSerIleAlaMetIysLeuIleIysLeuIysLeuGlnGlyTyrGlnLeu 80
 DB 358 CAAGTGGAGTTTCCATAGCTATGAACCTATCAAACTGAGCTACAGAGATATCAGCTA 417
 QY 81 ProSerAlaLeuProProValMet 88
 DB 418 CCCTCTGCACTTCCCTCTGCATG 441

RESULT 7
 AX880753 2131 bp DNA linear PAT 17-DEC-2003
 LOCUS Sequence 15658 from Patent EP1074617.
 AX880753
 AX880753.1 GI:40035489

SOURCE
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,D.,
 Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
 Primers for synthesizing full-length cDNA and their use
 Patent: EP 1074617-A 15658 07-FEB-2001.
 Research Association for Biotechnology (JRP)

TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 1. 2131
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 347..>2131
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAE92347.1"
 /db_xref="GI:40035490"
 /translation="WAQPTTPFGSLDTWATVEERAKHQPHSLKPIGSGITNGDA
 RNFPGSLPQPVLAQITWALDMNMDGMDOVERSIAMKLITKLQIGYOLPSALPPVM
 KOQPAISSAPFAVAPPLANGAPVIOPLAFAPHPATLTKSSFSRSGSQNLTK
 LOKASPDVASVPVAEMAVPOSRLKYOLFNSHDKMTSGHLGPOARTILMOSJLP
 OAOIASIWNLSIDIDOGKLTAEFFILAMHLIDVMSGQPPVLPPIYIPSPFRVRS
 GSGISVTSISVSVDRLPBEPLDEOOLLEKKLPVTEDEKRENERGNELEERROA
 LLEORKEORLEAQLERAEOREREREROREROREROLEEROLEEROLEERORERERK
 EIERERAEKLEEROLEERORERERORERORERORERORERORERORERORERORER
 KKHOLEKLEADICRLITTOROLEISTNSKSLRLAETITHLQOOLQESQOMIGRLIPK
 QIINDOLKOVQOINSIRDSIVTLKRALFAEKALRQHLRDLQDEVEKEKTRSKLQEDIDIF
 NNOLKEIREIHNKQLOKOKSMEARLKQ"

ORIGIN
 Alignment Scores:
 Pred. No.: 7.67e-53 Length: 2131
 Score: 459.00 Matches: 88
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 6

US-09-720-934-2_copy_15_102 (1-88) x AX880753 (1-2131)
 QY 1 TTPAlaIleThrValGluGluArgAlaYshHsAspGlnGlnPheHisSerLeuIysPro 20
 DB 178 TGGGCATTAAGTCTAGAGGAAAGAGGAGCATGATCAGAGTTCCATAGTTAAAGCCA 237

DB 389 TGGGCATTAAGTCTAGAGGAAAGAGGAGCATGATCAGAGTTCCATAGTTAAAGCCA 448
 QY 21 IIsErglyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
 DB 449 ATATCTGGATTCTATCTGTGATCAAGCTAGAAACCTTTTTCATCTGGATTACCT 508
 QY 41 GlnProValLeuAlaGlnIleThrPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
 DB 509 CAACCTGTTTACCAAGATATGGGCACTAGCTGACATGATATATGATGGAAGAATGGAT 568
 QY 61 GlnValGluPheSerIleAlaMetIysLeuIleIysLeuIysLeuGlnGlyTyrGlnLeu 80
 DB 569 CAAGTGGAGTTTCCATAGCTATGAACCTATCAAACTGAGCTACAGAGATATCAGCTA 628
 QY 81 ProSerAlaLeuProProValMet 88
 DB 629 CCCTCTGCACTTCCCTCTGCATG 652

RESULT 8
 BD127168 2131 bp DNA linear PAT 18-SEP-2002
 LOCUS Primer for synthesizing full-length cDNA and use thereof.
 BD127168
 BD127168.1 GI:23222113
 JP 2002017375-A/2599.
 JP 2002017375-A/2599.

SOURCE
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
 Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
 Koga,H.
 Primers for synthesizing full-length cDNA and use thereof
 Patent: JP 2002017375-A 2599 22-JAN-2002;
 HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002017375-A/2599

TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 1. 2131
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 FT CDS
 Location/Qualifiers
 (205)..(2130).

ORIGIN
 Alignment Scores:
 Pred. No.: 7.67e-53 Length: 2131
 Score: 459.00 Matches: 88
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 6

US-09-720-934-2_copy_15_102 (1-88) x BD127168 (1-2131)
 QY 1 TTPAlaIleThrValGluGluArgAlaYshHsAspGlnGlnPheHisSerLeuIysPro 20
 DB 247 TGGGCATTAAGTCTAGAGGAAAGAGGAGCATGATCAGAGTTCCATAGTTAAAGCCA 306

QY 21 11eserGlyphe1lethrGlyaspGlnAlaAargAsnPhenheGlnserGlyLeuPro 40
 DB 307 AATATCGATTCTACTGATGATCAAGCTAGAACTTTTTCATCTGGGTTACT 366
 QY 41 GlnProValLeuAlaGlnIleTrrPalalaLeuAlaaspMetAsnAsnAspGlyArgMetAsp 60
 DB 367 CAACCTGTTTTCACAGATATGGGACATGAGCATGATATATATGAAAGAAAGGAT 426
 QY 61 GlnValGluPheser1leAlaMetLysLeu1leLysLeuLysLeuGlnGlyTyrGlnLeu 80
 DB 427 CAAGTGAGTTTCCATCTACTATGAACCTATCAACTCAAGCTACAGAGATATTCAGCTA 486
 QY 81 ProSerAlaLeuProProValMet 88
 DB 487 CCCTCGACCTCCCTCTCATG 510
 RESULT 9
 LOCUS BD158570 2131 bp DNA linear PART 17-JAN-2003
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION BD158570
 VERSION BD158570.1 GI:27864328
 KEYWORDS JP 2002191363-A/13413.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 2131)
 Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Makamatsu,A., Nagai,K. and Otsuki,T.
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 13413 09-JUL-2002;
 HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002191363-A/13413
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU PI
 JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI MAKAMATSU, KI
 KEICHI NAGAI, TETSUKI OTSUKI
 PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
 10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key
 Location/Qualifiers
 FT CDS (347)..(2131).
 FEATURES
 source 1..2131
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.67e-53 Length: 2131
 Score: 459.00 Matches: 88
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 DB: 6
 US-09-720-934-2_COPY_15_102 (1-88) x BD158570 (1-2131)
 QY 1 TrrPalalThrValGluGluArgAlaLysHisaspGlnGlnPheHisSerLeuLysPro 20
 DB 389 TGGGCCATTACTGTAGAGAAAGGCGAAGCATATACAGCATTCATAGTTTAAAGCCA 448
 QY 21 11eserGlyphe1lethrGlyaspGlnAlaAargAsnPhenheGlnserGlyLeuPro 40
 DB 449 AATATCGATTCTACTGATGATCAAGCTAGAAACTTTTTCATCTGGGTTACT 508

QY 41 GlnProValLeuAlaGlnIleTrrPalalaLeuAlaaspMetAsnAsnAspGlyArgMetAsp 60
 DB 509 CAACCTGTTTTCACAGATATGGGACATGAGCATGATATATATGAAAGAAAGGAT 568
 QY 61 GlnValGluPheser1leAlaMetLysLeu1leLysLeuLysLeuGlnGlyTyrGlnLeu 80
 DB 569 CAAGTGAGTTTCCATCTACTATGAACCTATCAACTCAAGCTACAGAGATATTCAGCTA 628
 QY 81 ProSerAlaLeuProProValMet 88
 DB 629 CCCTCGACCTCCCTCTCATG 652
 RESULT 10
 LOCUS AK027846 2131 bp mRNA linear PRI 01-AUG-2002
 DEFINITION Homo sapiens cDNA FLJ14940 fis, clone PLACE1010942, highly similar to Homo sapiens intersectin long isoform (ITSN) mRNA.
 ACCESSION AK027846
 VERSION AK027846.1 GI:14042823
 KEYWORDS oligo cloning; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1
 Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M., Hiraoka,S., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Makamatsu,A., Nakamura,Y., Nagahari,K., Masuko,Y. and Kanehori,K.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2131)
 Isogai,T. and Otsuki,T.
 Direct Submission
 Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-1975, Fax:81-438-52-3586)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
 FEATURES
 source 1..2131
 Location/Qualifiers
 1..2131
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE1010942"
 /tissue_type="placenta"
 /clone_id="PLACE1"
 /note="Cloning vector: pME18SFL3"
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.67e-53 Length: 2131
 Score: 459.00 Matches: 88
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 DB: 9
 US-09-720-934-2_COPY_15_102 (1-88) x AK027846 (1-2131)
 QY 1 TrrPalalThrValGluGluArgAlaLysHisaspGlnGlnPheHisSerLeuLysPro 20
 DB 389 TGGGCCATTACTGTAGAGAAAGGCGAAGCATATACAGCATTCATAGTTTAAAGCCA 448
 QY 21 11eserGlyphe1lethrGlyaspGlnAlaAargAsnPhenheGlnserGlyLeuPro 40
 DB 449 AATATCGATTCTACTGATGATCAAGCTAGAAACTTTTTCATCTGGGTTACT 508

QY 41 GlnProValIleuAlaGlnIleTyrPalAlaLeuAlaAspMetCAsnAsnAspGlyAraGMetAsp 60
 DB 509 CAACCTGTTTACACAGATATGCGACATGACATGATATATGATGAGAAATGAGAT 568
 QY 61 GlnValGlnPheSerIleAlaMetCysLeuIleIleValLeuIleValGlnGlyTyrGlnIleu 80
 DB 569 CAAGTGAAGTTTCCATGACTATGAACTTATCAAACTGAAGCTACAGATATACAGCTA 628
 QY 81 ProSerAlaIleuProProValMet 88
 DB 629 CCTCTGCACCTTCCCTCTGCATG 652

RESULT 11
 AK074554 2131 bp mRNA linear PRI 03-SEP-2002
 AK074554
 LOCUS
 DEFINITION
 Homo sapiens cDNA FLJ90073 fis, clone HEMBA1004110, highly similar
 to Homo sapiens intersectin short form mRNA.
 ACCESSION
 AK074554 GI:22760070
 VERSION
 AK074554.1 GI:22760070
 KEYWORDS
 oligo capping; fis (full insert sequence).
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 AUTHORS
 1
 Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T.,
 Sugiyama, T., Suiko, Y., Nagai, K., Sugano, S., Ishii, S.,
 Kawai-Hito, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,
 Kojima, S., Nagahara, K., Masuno, Y., Ono, T., Okano, K., Yoshikawa, Y.,
 Aotsuma, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and
 Minomura, K.
 NEDO human cDNA sequencing project
 TITLE
 Unpublished
 JOURNAL
 2 (bases 1 to 2131)
 REFERENCE
 Isogai, T. and Otsuki, T.
 DIRECT SUBMISSION
 Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
 Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 COMMENT
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA library construction;
 Research Association for Biotechnology; cDNA library construction;
 Institute of Medical Science, University of Tokyo, Laboratory of
 Genome Structure, Human Genome Center; cDNA 5' - & 3' end one pass
 sequencing and clone selection; Helix Research Institute (supported
 by Japan Key Technology Center etc.).

FEATURES
 SOURCE
 1..2131
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="HEMBA1004110"
 /issue_type="whole embryo, mainly head"
 /clone_lib="HEMBA1"
 /dev_stage="embryo, 10 weeks"
 /note="cloning vector: PME18SFL3"

ORIGIN
 Alignment Scores: Length: 2131
 Pred. No.: 7.67e-53
 Score: 459.00 Matches: 88
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 9

US-09-720-934-2_COPY_15_102 (1-88) x AK074554 (1-2131)

QY 1 TTPAlaIleThrValGlnGlnArgAlaLysHisAspGlnPheHisSerIleuLysPro 20
 DB 247 TGGGCGATTAAGCTGTAAGAGAAAGAGGAAAGCATGATCAGCAGTTCATATGTTTAAAGCA 306

QY 21 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
 DB 307 ATATCTGATTCATTACTGCTGATCATCAGCTAGAACTTTTTCATCTGGGTTACCT 366
 QY 41 GlnProValIleuAlaGlnIleTyrPalAlaLeuAlaAspMetCAsnAsnAspGlyAraGMetAsp 60
 DB 367 CAACCTGTTTACACAGATATGCGACATGACATGATATATGATGAGAAATGAGAT 426
 QY 61 GlnValGlnPheSerIleAlaMetCysLeuIleIleValLeuIleValGlnGlyTyrGlnIleu 80
 DB 427 CAAGTGAAGTTTCCATGACTATGAACTTATCAAACTGAAGCTACAGATATACAGCTA 486
 QY 81 ProSerAlaIleuProProValMet 88
 DB 487 CCTCTGCACCTTCCCTCTGCATG 510

RESULT 12
 BC058925 2199 bp mRNA linear PRI 05-NOV-2003
 BC058925
 LOCUS
 DEFINITION
 Homo sapiens intersectin 1 (SH3 domain protein), mRNA (cDNA clone
 IMAGE:4443129), partial cds.
 ACCESSION
 BC058925 GI:37589134
 VERSION
 BC058925
 KEYWORDS
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 AUTHORS
 1 (bases 1 to 2199)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, J., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.D., Malek, U.A., Gunaratne, P.H., Richards, S.,
 Wooley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Bulky, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butcherfield, Y.S., Krzywinski, M.I., Skalska, U., Smilins, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
 MEDLINE
 22388257
 12477932
 2 (bases 1 to 2199)
 REFERENCE
 Strausberg, R.
 DIRECT SUBMISSION
 Submitted (01-OCT-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butcherfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Leticia Hsiao, Martin Krzywinski, Keta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeay, Steven
 Nese, Pawan Pandoh, Anna-Ilisa Prabh, Parvaneh Saeedi, Jacqueline

3

3

```

QY      61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
      461 CAAGTGGAGTTTCCATAGCTATGAACTTATCAAACTAGCAAGATATACAGCTA 520
DB      81 ProSerAlaLeuProProValMet 88
      521 CCTCTGCACTTCCCTCTGTCATG 544

RESULT 14
BD205033
LOCUS    BD205033
DEFINITION Isolated SH3 gene relating to myeloproliferative disorders and
            leukemia and utilization thereof.
ACCESSION BD205033
VERSION   BD205033.1 GI:33014803
KEYWORDS  JP 2002511267-A/1.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 5199)
AUTHORS   Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL   Korenberg,J.R. and Chen,X.N.
          Isolated SH3 gene relating to myeloproliferative disorders and
          leukemia and utilization thereof
          Patent: JP 2002511267-A 1 16-APR-2002;
          CEDARS SINAI HEALTH SYSTEM ET AL
          OS Homo sapiens (human)
          PN JP 2002511267-A/1
          PD 16-APR-2002
          PF 16-APR-1999 JP 2000543610
          PR 16-APR-1998 US 60/082007
          PI JULIE R KORENBERG,XIMO NING CHEN
          PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,
          PC C1201/68.
          PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
          CC Isolated SH3 gene relating to myeloproliferative disorders and
          CC leukemia
          CC and utilization thereof.
          FH Key location/Qualifiers
          FT source 1..5199
          FT location/Qualifiers
          FT source 1..5199
          FT /organism="Homo sapiens (human)".
          FT /db_xref="taxon:9606"

FEATURES
source      Location/Qualifiers
             1..5199
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 2.18e-52 Length: 5199
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-720-934-2_copy_15_102 (1-88) x BD205033 (1-5199)
QY      1 TTPAlaIleTTPValGlnGlnAlaArgAlaLysHISaSPGInGlnPheHISerLeuLysPro 20
      250 TGGGGCATACACTGAGAGAGAAAGACGAGCATGATGAGCATTCATGATTTAAAGCCA 309
DB      21 IlleSerGlyPheIleThrGlyAspGlnAlaArgaenPhePheGlnSerGlyLeuPro 40
      310 ATATTCGATTCATCTACTGGTCATCAAGCTAGAACTTTTTCATCTGGGTTACT 369
QY      41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetasp 60
      370 CAACCTGTTTTCACACAGATATGGGACATAGCTGACATGAATAAGATGAGAAATGAT 429
DB      61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80

```

```

DB      430 CAAGTGGAGTTTCCATAGCTATGAACTTATCAAACTAGCAAGATATACAGCTA 489
QY      81 ProSerAlaLeuProProValMet 88
      490 CCTCTGCACTTCCCTCTGTCATG 513

RESULT 15
AF064243
LOCUS    AF064243
DEFINITION Homo sapiens intersecin short form mRNA, complete cds.
ACCESSION AF064243
VERSION   AF064243.1 GI:3859852
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 5287)
AUTHORS   Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL   Guipponi,M., Scott,H.S., Chen,H., Schebesta,A., Rossier,C. and
          Anttonarakis,S.E.
          Two isoforms of a human intersecin (ITSN) protein are produced by
          brain-specific alternative splicing in a stop codon
          Genomics 53 (3), 369-376 (1998)
          MEDLINE 99017974
          PUBMED 9799604
          REFERENCE 2 (bases 1 to 5287)
          AUTHORS   Guipponi,M., Scott,H.S., Chen,H., Schebesta,A., Rossier,C. and
          Anttonarakis,S.E.
          Direct Submission
          Submitted (05-MAY-1998) Genetics and Microbiology, CMU, 1 rue
          Michel-Servet, Geneva 4 CH-1211, Switzerland
          Location/Qualifiers
          source      1..5287
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.1-q22.2"
                     /tissue_type="brain"
                     /dev_stage="fetus"
                     107..3769
                     /codon_start=1
                     /product="intersecin short form"
                     /protein_id="AAC78610.1"
                     /db_xref="GI:3859853"
                     /translation="MAQPTTFPGSGSLDIWAIIVERRAKHDQPHSLKPIGFTTGDQA
                     RNFFFGSLQPPVLAQIWMALDMNNDKMDQVERSIIMKLIKLGQQLPSALPVM
                     KOOPVAISSAPPPGMGIIASMPPLTAAVPMGSIPIVAGMSPPTLVSSVPTAAVPLAN
                     GAPVIOPLPAPAPHPATLPRKSSFSRSGSGSOLNTLOKAOSDFVAVPAAVAVP
                     OSSRLKRLQLFNSHDKTMSGHLTGPOARKTIMOSLPOAQLASITWISDIDQDKLTA
                     BEFTLAHNLIDVANSQGPLPVLPVLPPTIPSPFRVRSRGSGISVTSSTVDORLPEBPV
                     LEEDQQLLEKLLPVTFEDKRENERENLELEKSRQALLBQQRKEQERLADLEAEQ
                     RKREEROERKQLELEKQLEKRELEROEERREIERERAKRELEROLEQLE
                     RNREOELENGKOEQEDIVLKKAKKLELELELELELELELELELELELELELELELELE
                     EIESTNSRELELEITHLOQLOESOOMGLRILPEKQILNDQKVOQNSLHDSLV
                     TLKRLAKRELELEITHLOQLOESOOMGLRILPEKQILNDQKVOQNSLHDSLV
                     MEERLROKEDERKITELEKKEAEQRRARQDRDQWLEHVGQREHQRPRKLEHEKX
                     KREESVKKKQDEERKQEAQDKLRLTHQDEPAKPAVQAEMSTRKEKPLITSQENV
                     KVVYRALVPFESRSHDEITLQPDIVVWKEWDESGTGGWLGGLKRTGTFPA
                     NYAEKIPENEVAPKVTDSISAPAKLARLETPALATGSSSPPTNNWAFSST
                     MPSTNKEPETDNMDAMAQPSLTVPSAGQURQSAFPTATATGSSSPVVGQGXVE
                     GLOQALYPMRKAQDNHLENKKNVITVLEQDMMWGEVVGQGWGPFKSVKLIQSP
                     IRKSTSDSGSSSEKSPASIKRYASPAKRVYSGEFTIATWYTESSECGGLTPOCGVIL
                     VTAKDQDMWTSTVDKGVFSPNSVYRLKDSGSGTAGTGLGKKPELAQVITATAT
                     GPBQITLAPGQLILIRKKNPGWMEGELQAGKRRQQLGPPANVYKLNPGTSKITPT
                     EPPQSTALAAVCQVIGMYDYTAQNDDELAFNKGOIINVINKEDPDMWKEVNGVGLF
                     PSNVYKLTITMDPESQQ"
          misc_feature 167..406
                     /note="encodes EH domain"
          misc_feature 767..1936
                     /note="encodes EH domain"
          misc_feature 2324..2524

```

```

misc_feature      /note="encodes SH3 domain"
                  2843..3019
                  /note="encodes SH3 domain"
misc_feature      3110..3286
                  /note="encodes SH3 domain"
misc_feature      3326..3520
                  /note="encodes SH3 domain"
misc_feature      3569..3748
                  /note="encodes SH3 domain"
ORIGIN

```

```

Alignment Scores:
Pred. No.:      2.22e-52
Score:          459.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match:    100.00%
DB:             9

```

US-09-720-934-2_COPY_15_102 (1-88) * AF064243 (1-5287)

```

QY      1 TrpAlaIleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
          |||
DB      149 TGGGCCATTAAGAGGAAAGAGGAGCATGATCAGCAGTCCATAGTTAAAGCCA 208
          |||
QY      21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
          |||
DB      209 ATATCTGGATTCATTACTGCTGATCAAGCTAGAACTTTTTCATCTGGTTACCT 268
          |||
QY      41 GlnProValLeuAlaGlnIleThrPalaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
          |||
DB      269 CAACCTGTTTTCAGCAGATATGGGCACTGACATGATAATATATGATGAGAGATGGAT 328
          |||
QY      61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeu 80
          |||
DB      329 CAAAGTGGAGTTTCCATAGCTTGAAGACTTATCAAACTGAAAGCTACAAAGATATCAGCTA 388
          |||
QY      81 ProSerAlaLeuProProValMet 88
          |||
DB      389 CCCTCTGCACCTCCCTGTCATG 412
          |||

```

Search completed: July 1, 2004, 16:14:04
Job time : 2015.31 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:01:26 ; Search time 2647.21 Seconds
(without alignments)
992.694 Million cell updates/sec

Title: US-09-720-934-2_COPY_15_102
Perfect score: 459
Sequence: 1 WAITEERAKHQDFHSIKP.....KLKIKLQGYQLPSALPPYM 88

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DRV=xlp
-O/cn2 1/USPTO.spool.p/US09720934/runat.30062004.064540.13442/app.query.fasta_1.1386
-DB-RST -QFMT=fastcap -SUPFIX=ret -MINMATCH=0.1 -LOOPEC=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US09720934 @CN2 1.1 12421 @runat.30062004.064540.13442 -NCPU=6 -ICPU=3
-NO MMAP -LARGEJOURNEY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEF_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthm:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rdd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	459	100.0	477	14	CB269094
2	459	100.0	544	9	AL7111737
3	459	100.0	547	10	BT194430
4	459	100.0	734	14	CB269754
5	459	100.0	896	13	BQ941336
6	459	100.0	979	13	BQ941411
7	459	100.0	1089	12	BQ050397
8	459	100.0	2079	11	BC013578
9	459	100.0	2126	11	BC020269
10	451	98.3	523	13	B0381021
11	451	98.3	907	13	B0128229
12	441	96.1	445	9	AI852079
13	441	96.1	475	10	BE199657
14	441	96.1	488	9	AI481705
15	441	96.1	501	9	AI852070
16	441	96.1	510	10	BE952536
17	441	96.1	533	29	CQ579249
18	441	96.1	600	14	CA528645
19	441	96.1	662	14	CF536420
20	441	96.1	698	14	CB519152
21	441	96.1	728	13	BY731854
22	441	96.1	728	13	CF729634
23	441	96.1	728	13	CF729634
24	441	96.1	728	13	CF729634
25	440	95.9	811	9	AI549192
26	440	95.9	811	12	BG829540
27	432	94.1	750	12	BG118422
28	431	93.9	643	13	BK470886
29	426	92.8	766	14	CF743758
30	423	92.2	723	10	BE786696
31	420	91.5	624	13	BY713771
32	410	89.3	359	9	AI853008
33	410	89.3	589	9	AL792490
34	410	89.3	822	9	AL870708
35	409	89.1	534	12	BG160235
36	409	89.1	550	12	B0623100
37	400	87.1	504	9	AL964682
38	365	79.5	422	14	CA388981
39	365	79.5	698	14	CA391925
40	359	78.2	399	9	AA217338
41	359	78.2	404	9	AI390576
42	359	78.2	490	9	AI646131
43	359	78.2	556	12	BG144398
44	359	78.2	565	9	AI809238
45	359	78.2	619	10	BB614005

ALIGNMENTS

RESULT 1
LOCUS CB269094 477 bp mRNA linear EST 20-FEB-2003
DEFINITION 1008001 Human Fat Cell 5'-Stretch plus cDNA Library Homo sapiens
ACCESSION CB269094
VERSION CB269094.1 GI:28443679
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 477)

AUTHORS Yang, R.-Z., Shuldiner, A. and Gong, D.-W.
 TITLE EST analysis of human adipose gene expression
 JOURNAL Unpublished (2002)
 COMMENT

Contact: Gong Da-Wei
 Division of Endocrinology, Diabetes and Nutrition
 University of Maryland
 660 Redwood St., HH497, Baltimore, MD 21201, USA
 Tel: 410 706 1672
 Fax: 410 706 1622
 Email: dgong@medicine.umaryland.edu

PCR Primers
 FORWARD: CTCGGGAAACCGCCGATTGTGTGTG
 BACKWARD: AATACGACTCATTAGGCGCAATTGG
 Seq primer: GTTGGTACCCGGAATTC.

FEATURES
 source Location/Qualifiers
 1..477

ORIGIN
 Alignment Scores:
 Pred. No.: 1,678-53 Length: 477
 Score: 459.00 Matches: 88
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 14 Gaps: 0

US-09-720-934-2_copy_15_102 (1-88) x CB269094 (1-477)

QY 1 TTPALAIETHRVALGUGUARGALALYSHISASPGINGINPHEHISERILEUYSRPRO 20
 Db 4 TGGGCGCATTAACGTAGAGAGAAAGAGCGAAGCATGATCGACATTCATAGTTAAAGCCA 63
 QY 21 ILESERGIYPHEIIEIETHRGILYASPGINLAARGANPHEPHEGINSERGIYLEUPRO 40
 Db 64 ATATCTGATTCATTAACGTGATCAAGTAACTTTTTCATCTGGGTTTACCT 123
 QY 41 GINPROVALLEUALAGNIIETTPALALEUALAASPMEASASRPGIYARGMECASP 60
 Db 124 CAACCTGTTTAGACACACATATGGCACTAGCTGACATGATTAATGATGAAGAATGAT 183
 QY 61 GINVALGUPHESERILEALAMELISLEULEISLEULEISLEUNGILYTYRGINLEU 80
 Db 184 CAAGTGGAGTTTCCATGACTATGAACTATCAAACTGAAGCTACAAAGATATCAGCTA 243
 QY 81 PROSERIALAENUPROPROVALMET 88
 Db 244 CCTCTGCACCTTCCCTCTGCATG 267

RESULT 2
 LOCUS AL711737 544 bp mRNA linear EST 04-SEP-2003
 DEFINITION DKFZP686K1884_r1_686 (synonym: hicc3) Homo sapiens cDNA clone
 ACCESSION DKFZP686K1884_5, mRNA sequence.
 VERSION AL711737
 KEYWORDS AL711737.1 GI:19695092
 SOURCE EST.

ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCES
 1 (bases 1 to 544)
 Duesterhoeft, A., Lauber, J., Mewes, H. W., Gaassenhuber, J. and
 Wiemann, S.
 EST (Duesterhoeft, et al.)
 TITLE JOURNAL
 COMMENT Unpublished (1999)
 Contact: MIPS

MIPS
 Ingolstaedter Landstr.1, D-85764 Neuberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 No sl sequence available.
 This clone (DKFZP686K1884) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
 source Location/Qualifiers
 1..544

ORIGIN
 Alignment Scores:
 Pred. No.: 1,986-53 Length: 544
 Score: 459.00 Matches: 88
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x AL711737 (1-544)

QY 1 TTPALAIETHRVALGUGUARGALALYSHISASPGINGINPHEHISERILEUYSRPRO 20
 Db 216 TGGGCGCATTAACGTAGAGAGAAAGAGCGAAGCATGATCGACATTCATAGTTAAAGCCA 275
 QY 21 ILESERGIYPHEIIEIETHRGILYASPGINLAARGANPHEPHEGINSERGIYLEUPRO 40
 Db 276 ATATCTGATTCATTAACGTGATCAAGTAACTTTTTCATCTGGGTTTACCT 335
 QY 41 GINPROVALLEUALAGNIIETTPALALEUALAASPMEASASRPGIYARGMECASP 60
 Db 336 CAACCTGTTTAGACACACATATGGCACTAGCTGACATGATTAATGATGAAGAATGAT 395
 QY 61 GINVALGUPHESERILEALAMELISLEULEISLEULEISLEUNGILYTYRGINLEU 80
 Db 396 CAAGTGGAGTTTCCATGACTATGAACTATCAAACTGAAGCTACAAAGATATCAGCTA 455
 QY 81 PROSERIALAENUPROPROVALMET 88
 Db 456 CCTCTGCACCTTCCCTCTGCATG 479

RESULT 3
 LOCUS BF194430 547 bp mRNA linear EST 02-NOV-2000
 DEFINITION 246586 MARRC 2P1G Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION BF194430
 VERSION BF194430.1 GI:11077799
 KEYWORDS EST.

ORGANISM Sus scrofa (pig)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 REFERENCES
 1 (bases 1 to 547)
 Fahrnenkung, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,
 Vallet, J., Wise, T., Rohrer, G.A., Pettes, G., Sultana, R.,
 Quackenbush, J. and Keele, J.W.
 Porcine gene discovery by normalized cDNA-library sequencing and
 EST cluster assembly
 Mamm. Genome 13 (8), 475-478 (2002)
 TITLE JOURNAL
 MEDLINE 22213789

PUBMED 12226715

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred

v0.960904.e. Vector identified by cross_match with the -minscore 18

and -mismatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCGCCATCAGCAGC

Plate: 76 Row: E Column: 21

Seq primer: ATTGAGTGACACTATAG.

Location/Qualifiers

1..547

/organism="Sus scrofa"

/mol_type="mRNA"

/db_xref="taxon:9823"

/tissue_type="pooled"

/lab_host="DH10B"

/clone_lib="MARC 2P1G"

/note="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI;

Library made from pooled tissue from testis, ovary, "

endometrium, hypothalamus, pituitary, and placenta."

ORIGIN

Alignment Scores:

Pred. No.:	1,996-53	Length:	547
Score:	459.00	Matches:	88
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
		Gaps:	0

US-09-720-934-2_COPY_15_102 (1-88) x BF194430 (1-547)

```
QY 1 TTPAalleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
Db 185 TGGGCCATTAAGTCTAGAGAAAGAGCAAGCATGATCAGAGTTCCATAGCTTAAAGCCA 244
QY 21 TleserGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnInseryLysPro 40
Db 245 ATATCTGATTTATTTACTGCGATCAAGCTAAGAACTTTTTCATCTGGGGTTAACCCT 304
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAspGlyArgMetAsp 60
Db 305 CAACCTGTTTTCATAGACAGATATGGCGCTGCTGATGATATATGACGGAAGATGAT 364
QY 61 GlnValGlnPheSerIleAlaMetLysLeuLysLeuLysLeuGlnGlyTyrGlnLeu 80
Db 365 CAAGTGGAGTTTTCATAGCAGATGAAGAACTTCAAGCTCAAGCTCAAGCTATCAGCTC 424
QY 81 ProSerAlaLeuProProValMet 88
Db 425 CCTCTCGCACTTCCCTCTGTCATG 448

RESULT 4
LOCUS CB269754 734 bp mRNA linear EST 20-FEB-2003
DEFINITION 1008661 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens
ACCESSION CB269754
VERSION CB269754.1 GI:28444339
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 734)
AUTHORS Yang,R.-Z., Shuldiner,A. and Gong,D.-W.
TITLE EST analysis of human adipose gene expression
```

JOURNAL

Unpublished (2002)

Contact: Gong Da-Wei

Division of Endocrinology, Diabetes and Nutrition

University of Maryland

660 Redwood St, HH497, Baltimore, MD 21201, USA

Tel: 410 706 1672

Fax: 410 706 1622

Email: dgong@medicine.umaryland.edu

PCR Primers

FORWARD: CTCGGAGAGCGCCATTTGTGTGT

BACKWARD: AATAGACTACTATAGAGCGAATTC

Seq primer: GTTGTATCCCGGAATTC.

Location/Qualifiers

1..734

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/sex="Male and Female"

/tissue_type="Adipose"

/clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"

/note="Vector: lambdaBaltip1Bx"

ORIGIN

Alignment Scores:

Pred. No.:	2,916-53	Length:	734
Score:	459.00	Matches:	88
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
		Gaps:	0

US-09-720-934-2_COPY_15_102 (1-88) x CB269754 (1-734)

```
QY 1 TTPAalleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
Db 163 TGGGCCATTAAGTCTAGAGAAAGAGCAAGCATGATCAGAGTTCCATAGCTTAAAGCCA 222
QY 21 TleserGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnInseryLysPro 40
Db 223 ATATCTGATTTATTTACTGCGATCAAGCTAAGAACTTTTTCATCTGGGGTTAACCCT 282
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAspGlyArgMetAsp 60
Db 283 CAACCTGTTTTCATAGACAGATATGGCGCTGCTGATGATATATGACGGAAGATGAT 342
QY 61 GlnValGlnPheSerIleAlaMetLysLeuLysLeuLysLeuGlnGlyTyrGlnLeu 80
Db 343 CAAGTGGAGTTTTCATAGCAGATGAAGAACTTCAAGCTCAAGCTCAAGCTATCAGCTA 402
QY 81 ProSerAlaLeuProProValMet 88
Db 403 CCTCTCGCACTTCCCTCTGTCATG 426

RESULT 5
LOCUS BQ941336 896 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT_8741326 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6420600
ACCESSION BQ941336
VERSION BQ941336.1 GI:22356814
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
CDNA Library Preparation: Rubin Laboratory
```

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM2594 row: k column: 01
High quality sequence stop: 763.

FEATURES

source

1..896
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:420600"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_18"
/note="Organ: lung; Vector: pOT7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 3,766-53 Length: 896
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BQ941336 (1-896)

QY 1 TTPAIAIETHRVAGLUGUARGALAYSHASPGINGINPHEHISERLEUYSPro 20
DB 252 TGGGCCATTAACCTGTAAGAGAAAGCGAAGCATGATCGACATTCATGTTAAAGCCA 311
QY 21 ILESERGIYPHEIETHRGIVASPGINAAARGAENPHEPHEGINSERGILEUPro 40
DB 312 ATATCTGATTCATTAACCTGTAAGAGAAAGCATGATTCATGTTAAAGCCA 371
QY 41 GINPROVALLLEUALAGINLETTPALALEUALAASPMECASNASPGIYARGMETASP 60
DB 372 CAACCTGTTTAGACACATATGCGCACTAGCTGACATGATTAAGATGAAGATTGAT 431
QY 61 GINVAIGUPHESERILEALAMETLYSLEULIETLYSLEULINGLYTYRGINDEN 80
DB 432 CAAGTGGAGTTTCCATAGCTATGAAACTTATCAAACTGAAAGCTACAAAGATATCAGCTA 491
QY 81 PROSERIALAENUPROPROVALMET 88
DB 492 CCTCTGCACTTCCCTGTCATG 515
RESULT 6
BQ941411
LOCUS BQ941411 979 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT 8744157 lupski_sciatic_nerve Homo sapiens cDNA clone
IMAGE:6205905 5', mRNA sequence.
ACCESSION BQ941411
VERSION BQ941411.1 GI:22356889
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cga@phs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM13629 row: i column: 10
High quality sequence stop: 599.

FEATURES

source

1..979
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6205905"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="lupski_sciatic_nerve"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; CDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TGACCCACGAGCGCG-3' and
5'-GACTAGTTCATGATGCGAGCGCGCCCTT(5)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

ORIGIN

Alignment Scores:

Pred. No.: 4,226-53 Length: 979
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BQ941411 (1-979)

QY 1 TTPAIAIETHRVAGLUGUARGALAYSHASPGINGINPHEHISERLEUYSPro 20
DB 328 TGGGCCATTAACCTGTAAGAGAAAGCGAAGCATGATCGACATTCATGTTAAAGCCA 387
QY 21 ILESERGIYPHEIETHRGIVASPGINAAARGAENPHEPHEGINSERGILEUPro 40
DB 388 ATATCTGATTCATTAACCTGTAAGAGAAAGCATGATTCATGTTAAAGCCA 447
QY 41 GINPROVALLLEUALAGINLETTPALALEUALAASPMECASNASPGIYARGMETASP 60
DB 448 CAACCTGTTTAGACACATATGCGCACTAGCTGACATGATTAAGATGAAGATTGAT 507
QY 61 GINVAIGUPHESERILEALAMETLYSLEULIETLYSLEULINGLYTYRGINDEN 80
DB 508 CAAGTGGAGTTTCCATAGCTATGAAACTTATCAAACTGAAAGCTACAAAGATATCAGCTA 567
QY 81 PROSERIALAENUPROPROVALMET 88
DB 568 CCTCTGCACTTCCCTGTCATG 591
RESULT 7
BQ050397
LOCUS BQ050397 1089 bp mRNA linear EST 29-MAR-2002
DEFINITION AGENCOURT 7050803 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5784342
5', mRNA sequence.
ACCESSION BQ050397
VERSION BQ050397.1 GI:19809737
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens


```

VERSION      BC020269.1  GI:17939664
KEYWORDS     HTC.
SOURCE       Homo sapiens (human)
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE    1 (bases 1 to 2126)
AUTHORS      Strausberg, R.
TITLE        Direct Submission
JOURNAL      Submitted (19-DEC-2001) National Institutes of Health, Mammalian
              Gene Collection (MGC), Cancer Genomics Office, National Cancer
              Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
              USA
REMARK       NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
              Email: cgabs-remail.nih.gov
              Tissue Procurement: ATCC
              CDNA Library Preparation: Rubin Laboratory
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Genome Sequence Centre,
              BC Cancer Agency, Vancouver, BC, Canada
              info@bcsc.bc.ca
              Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
              Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
              Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
              Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
              Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
              Schein, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Stott,
              Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
              George Yang, Scott Zydziedrzyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAL Plate: 40 Row: n Column: 1
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4504796
This clone has the following problem: no cloning site /
microdeletion.

FEATURES
    source         1..2126
                    location/Qualifiers
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:4899011"
                    /tissue_type="Pancreas, epithelioid carcinoma"
                    /clone_id="NTH MGC 42"
                    /lab_host="DH10B-R"
                    /note="Vector: pOTB7"

ALIGNMENT Scores:
Pred. No.:      1-156-52
Score:          459.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match:    100.00%
DB:             11
              Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BC020269 (1-2126)
OY      1 TTPAIAIETHRVALGUGUARGALALYSHISAASPGINGINPHEHISEXTLEUYSFRO 20
Db      224 TGGGCGATTAACCTGTAAGGAAAGAGCGAAGCATGATGACGATTCATAGTTTAAAGCA 283
OY      21 TLESERGILYPhEIIETHGILYASPGINLAARGAENPHEPNEGINSERGIYLEUDRO 40
Db      284 ATATCTGATTCATTAACGTGATCAAGCTAGAAAACITTTTTTCAATCTGGGTTAACC 343
OY      41 GINPROVALLEUALAGINIIETTPALALEUALAASPMECAASNASASPGIYARMEVESP 60
Db      344 CAACCTGTTTTCAGACAGATATGGGCACTAGCTGACATGAATATGATGAGAAAGATGGAT 403
OY      61 GINVALGUPHESERILEALAMECLYSLEULEIYLSLEUYSLEUGINGLYTYRGINDUEN 80

```

```

Db      404 CAACTGAGCTTTTCCACTAGCTATGAAACTTATCAAACTAGAGCTACAGGATATGCTA 463
OY      81 PROSERIALALEUPROPROVALMET 88
Db      464 CCTCTGCACTTCCCTCTGTCATG 487

RESULT 10
BU381021
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Simon Hubbard
Department of Biomedical Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 0161360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
    source         1..523
                    location/Qualifiers
                    /organism="Gallus gallus"
                    /mol_type="mRNA"
                    /strain="White Leghorn, HiseX"
                    /db_xref="taxon:9031"
                    /clone="CHEST872a15"
                    /dev_stage="36"
                    /lab_host="DH10B"
                    /clone_id="CSEQCHN75"
                    /note="Organ: trunks; Vector: pBluescript II KS(+);
                    Site 1: EcoRI; Site 2: NotI; This normalized library was
                    constructed from 1 million independent clones. cDNA
                    synthesis was initiated using an oligo(dT) primer, using
                    methylated C in the first strand synthesis reaction.
                    Following this first strand reaction, double-stranded cDNA
                    was blunted, ligated to NotI adapters, digested with
                    EcoRI, size-selected, and cloned into the NotI and EcoRI
                    compatible sites of a custom modified MCS of the
                    pBluescript (KS+) vector. The library was normalized in 2
                    rounds using conditions adapted from Soares et al., PNAS
                    (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
                    (1996) 791, except that a significantly longer
                    reannealing hybridization was used."

ORIGIN
ALIGNMENT Scores:
Pred. No.:      2-456-52
Score:          451.00
Percent Similarity: 98.86%
Best Local Similarity: 97.73%
Query Match:    98.26%
DB:             13
              Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BU381021 (1-523)
OY      1 TTPAIAIETHRVALGUGUARGALALYSHISAASPGINGINPHEHISEXTLEUYSFRO 20

```

Db 196 TGGGCTATTACTGTGAGAGAGAGCTAAACATGATCAACAGTTCCACAGTCTGAAACCA 255

QY 21 TlSerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40

Db 256 ACATCTGAGATTTCATCTGCTGATCAAGCTAGAAACTTTTTCAGTCTGGGTTAACTT 315

QY 41 GlnProValLeuAlaGlnIleTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60

Db 316 CAGCCGGTGTAGACAGATATGGGCTTTAGCTGACATGAACATGATGAGAGATGGAT 375

QY 61 GlnValGlnPheSerIleAlaMetIleLeuIleLysLeuLysLeuGlnGlyTYrGlnLeu 80

Db 376 CAGCTGAGATTTCATAGCTATGAACTTATCAAAATTAAACTACAGAGCTATCACTC 435

QY 81 ProSerAlaLeuProProValMet 88

Db 436 CCATCTGCTGCTGCTCTGTCATG 459

RESULT 11

LOCUS BU128229 907 bp mRNA linear EST 25-NOV-2002

DEFINITION 603114019F1 CSECHL20 Gallus gallus CDNA clone CHEST6566 5', mRNA

ACCESSION BU128229

VERSION BU128229

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 907)

AUTHORS Boardman, P.E., Sans-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, M.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.C.

TITLE A Comprehensive Collection of Chicken CDNAs

JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)

MEDLINE 22335534

PUBMED 12445392

COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers

FEATURES

source

1..907

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="White Leghorn, Hisex"

/db_xref="taxon:9031"

/clone="CHEST6566"

/dev_stage="36"

/lab_host="DH10B"

/clone_lib="CSECHL20"

/note="Organ: Limbs; Vector: pBluescript II KS(+); Site: 1: EcoRI; Site 2: NotI; Modification of pBluescript II KS(+) [Stratagene] vector to accommodate cDNA produced with the T-trimmed protocol (Construction of uni-directionally cloned cDNA libraries from messenger RNA for improved 3' end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI. Ligate in double stranded adaptor containing BspI and BamHI sites [5'ggccggctgcagcccgagatccgaaataag] [5'aattcttttcggtccgagcgagcgc]"

ORIGIN

Alignment Scores:

Pred. No.: 4.99e-52 Length: 907

Score: 451.00 Matches: 86

Percent Similarity: 98.86% Conservative: 1

Best Local Similarity: 97.73% Mismatches: 1

Query Match: 98.26% Indels: 0

DB: 13 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BU128229 (1-907)

QY 1 TlPAlleIleThrValGluGluArgAlaIleAspGlnGlnPheHisSerLeuIlePro 20

Db 62 TGGGCTATTACTGTGAGAGAGAGCTAAACATGATCAACAGTTCCACAGTCTGAAACCA 121

QY 21 TlSerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40

Db 122 ACATCTGAGATTTCATCTGCTGATCAAGCTAGAAACTTTTTCAGTCTGGGTTAACTT 181

QY 41 GlnProValLeuAlaGlnIleTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60

Db 182 CAGCCGGTGTAGACAGATATGGGCTTTAGCTGACATGAACATGATGAGAGATGGAT 241

QY 61 GlnValGlnPheSerIleAlaMetIleLeuIleLysLeuLysLeuGlnGlyTYrGlnLeu 80

Db 242 CAGCTGAGATTTCATAGCTATGAACTTATCAAAATTAAACTACAGAGCTATCACTC 301

QY 81 ProSerAlaLeuProProValMet 88

Db 302 CCATCTGCTGCTGCTCTGTCATG 325

RESULT 12

LOCUS AI852079 445 bp mRNA linear EST 15-JUL-1999

DEFINITION UI-M-BH0-aj-a-f-07-0-UI.s1 NIH BMAP M.S1 Mus musculus CDNA clone

ACCESSION AI852079

VERSION AI852079

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 445)

AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD 20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mst@mail.nih.gov
Oligo-dT track not found. Not 1 site shown in beginning of sequence is likely internal to the message. cDNA library preparation: M.B. Soares Lab Clone distribution: NIH BMAP cDNA clones will be made available by the means that is soon to be determined. When NIH determines the means for distribution of the BMAP cDNA clones, this record will be updated accordingly when that means is determined.
Seq primer: M13 Forward
POLYA=No.
Location/Qualifiers

FEATURES

source

1..445

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UI-M-BH0-aj-a-f-07-0-UI"

/dev_stage="27-32 days"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NIH BMAP M.S1"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH BMAP M.S1 library is a subtracted library derived from a mixture of normalized libraries from ten regions of the

mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, dorex, amygdala, basal ganglia, pineal gland, striatum, hippocampus). The driver used for subtraction consisted of a pool of 20,000 cDNA clones obtained from non-normalized and normalized libraries of these ten regions of the mouse brain.
TAG_TISSUE=basal-ganglia
TAG_LIB=NIH_BMAP_M_S1
TAG_SEQ=TGTAC"

ORIGIN

Alignment Scores:

Pred. No.: 4,93e-51 Length: 445
Score: 441.00 Matches: 85
Percent Similarity: 97.73% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 2
Query Match: 96.08% Indels: 0
De: 9 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x A1852079 (1-445)

QY 1 TTPAIAIETThrValGluGluArgAlaYshHisAspGlnGlnPheHisSerLeuYsPro 20
Db 103 TGGGCCATTAAGTGTGAGAGAAAGGCGCAAGCATGACACGACTTCTTAGCCTGAAGCCG 162
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
Db 163 ATAGCGGATTTATTAAGTGTATCAAGGAGAACTTTTTCATTCATCGGTTACT 222
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
Db 223 CAGCTGCTTTCAGCAATATATGCGCGCTAGCGGACATGAATACGATGAGATGAT 282
QY 61 GlnValGluPheSerIleAlaMetYsLeuIleYsLeuYsLeuGlnGlyTyrglnLeu 80
Db 283 CAAGTGAATTTTCCATGACCATGAGCTTATCAACTGAAGCTACAGATATACAGCTC 342
QY 81 ProSerAlaLeuProProValMet 88
Db 343 CCTCCACACTTCCCTGTCATG 366

RESULT 13

LOCUS BE199657 475 bp mRNA linear EST 26-JUN-2000
DEFINITION uc70a11.x1 Soares mammary gland NMLMG Mus musculus cDNA clone
IMAGE:1547708 3' similar to TR:Q9ZOR4 Q9ZOR4 ESEIL PROTEIN. ;, mRNA
sequence.

ACCESSION BE199657 GI:8711826

VERSION BE199657.1

KEYWORDS Mus musculus (house mouse)

SOURCE EST.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 475)
AUTHORS NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:953056
Possible reversed clone: similarity on wrong strand
High quality sequence stop: 469.
Location/Qualifiers

FEATURES
source 1..475
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:1547708"

/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/clone_lib="Soares mammary_gland_NMLMG"
/note="Vector: pT7T3-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT7T3 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Alignment Scores:
Pred. No.: 5,37e-51 Length: 475
Score: 441.00 Matches: 85
Percent Similarity: 97.73% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 2
Query Match: 96.08% Indels: 0
De: 10 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x BE199657 (1-475)

QY 1 TTPAIAIETThrValGluGluArgAlaYshHisAspGlnGlnPheHisSerLeuYsPro 20
Db 104 TGGGCCATTAAGTGTGAGAGAAAGGCGCAAGCATGACACGACTTCTTAGCCTGAAGCCG 163
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
Db 164 ATAGCGGATTTATTAAGTGTATCAAGGAGAACTTTTTCATTCATCGGTTACT 223
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
Db 224 CAGCTGCTTTCAGCAATATATGCGCGCTAGCGGACATGAATACGATGAGATGAT 283
QY 61 GlnValGluPheSerIleAlaMetYsLeuIleYsLeuYsLeuGlnGlyTyrglnLeu 80
Db 284 CAAGTGAATTTTCCATGACCATGAGCTTATCAACTGAAGCTACAGATATACAGCTC 343
QY 81 ProSerAlaLeuProProValMet 88
Db 344 CCTCCACACTTCCCTGTCATG 367

RESULT 14

LOCUS A1481705 488 bp mRNA linear EST 09-MAR-1999
DEFINITION wh30f01.x1 Barstead mouse pooled organs MPEB4 Mus musculus cDNA
clone IMAGE:888505 3' similar to TR:O42287 O42287 INTERSECTIV. ;,
mRNA sequence.

ACCESSION A1481705 GI:4374931

VERSION A1481705.1

KEYWORDS Mus musculus (house mouse)

SOURCE EST.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 488)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE Unpublished (1999)

JOURNAL Contact: Marra M/WashU-NCI Mouse EST Project 1999
COMMENT Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Fri Jul 2 10:23:37 2004

us-09-720-934-2_copy_15_102.rst

Page 10

Search completed: July 1, 2004, 19:48:26
Job time : 2655.21 secs


```

Db      116 TCAGAGGTTACATACAGAGTATCAAGCCGCTACTTTTCTCTACAGCTCTGCCG 175
QY      41 GlnProValLeuAlaGlnIleTrrPaLaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
Db      176 GCCCGGTTTACCTGGAATATGGGCTTATCAGATCTGAACAAGATGGGAAGATGGAC 235
QY      61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeu 80
Db      236 CAGCAAGAGTTCTCTATACCTATGAACCTCATCAAGTTAAAGTTGACAGGCCAAGCTG 295
QY      81 ProSerAlaLeuProProValMet 88
Db      296 CCTGAGTCTCTCCCTCATCATG 319

RESULT 2
US-09-338-933-60
; Sequence 60, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338,933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-60

Alignment Scores:
Pred. No.: 1.26e-46      Length: 480
Score: 359.00           Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21%      Indels: 0
                        Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-338-933-60 (1-480)
QY      1 TrpAlaIleThrValGlnGluArgAlaLysHISAspGlnGlnPheHISerLeuLysPro 20
Db      56 TGGGCTATTACATCTGAAGAAAGTACTAAGCATGATTAACAAGTTGATTAACCTCAAACT 115
QY      21 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
Db      116 TCAGAGGTTACATACAGAGTATCAAGCCGCTACTTTTCTCTACAGCTCTGCCG 175
QY      41 GlnProValLeuAlaGlnIleTrrPaLaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
Db      176 GCCCGGTTTACCTGGAATATGGGCTTATCAGATCTGAACAAGATGGGAAGATGGAC 235
QY      61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeu 80
Db      236 CAGCAAGAGTTCTCTATACCTATGAACCTCATCAAGTTAAAGTTGACAGGCCAAGCTG 295
QY      81 ProSerAlaLeuProProValMet 88
Db      296 CCTGAGTCTCTCCCTCATCATG 319

RESULT 3
US-09-215-681-60
; Sequence 60, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; APPLICANT: King, Gordon E.

```

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; FILE REFERENCE: 210121.463
; CURRENT APPLICATION NUMBER: US/09/215,681A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-215-681-60

Alignment Scores:
Pred. No.: 1.26e-46      Length: 480
Score: 359.00           Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21%      Indels: 0
                        Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-215-681-60 (1-480)
QY      1 TrpAlaIleThrValGlnGluArgAlaLysHISAspGlnGlnPheHISerLeuLysPro 20
Db      56 TGGGCTATTACATCTGAAGAAAGTACTAAGCATGATTAACAAGTTGATTAACCTCAAACT 115
QY      21 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
Db      116 TCAGAGGTTACATACAGAGTATCAAGCCGCTACTTTTCTCTACAGCTCTGCCG 175
QY      41 GlnProValLeuAlaGlnIleTrrPaLaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
Db      176 GCCCGGTTTACCTGGAATATGGGCTTATCAGATCTGAACAAGATGGGAAGATGGAC 235
QY      61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeu 80
Db      236 CAGCAAGAGTTCTCTATACCTATGAACCTCATCAAGTTAAAGTTGACAGGCCAAGCTG 295
QY      81 ProSerAlaLeuProProValMet 88
Db      296 CCTGAGTCTCTCCCTCATCATG 319

RESULT 4
US-09-216-003A-60
; Sequence 60, Application US/09216003A
; Patent No. 6670463
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
; FILE REFERENCE: 210121.462
; CURRENT APPLICATION NUMBER: US/09/216,003A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-216-003A-60

Alignment Scores:
Pred. No.: 1.26e-46      Length: 480
Score: 359.00           Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21%      Indels: 0
                        Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-216-003A-60 (1-480)

```



```
QY 1 TTPAlaIleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
    |||||
Db 56 TGGGCTATTACATGTAAGAACGACTAAGCATGATTAACAAGTTGATTAACCTCAAACT 115
QY 21 ILeSerGlyPheIleThrGlyAspGlnAlaArgsnPhePheGlnSerGlyLeuPro 40
    |||||
Db 116 TCAGAGGTATACATAAAGGTGATCAAGCCCGTACTTTTCCACAGCTCAGGCTCTCCG 175
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetCsnAsnAspGlyArgMetAsp 60
    |||||
Db 176 GCCCGGTTTAGCTGAATAATGGGCTTATCAGATCTGAACAGAGATGGGAGATGTGAC 235
QY 61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
    |||||
Db 236 CAGCAAGAGTTCTTATAGCTATGAACTCATCAAGTTAAAGTTGCAAGGCCAACAGCTG 295
QY 81 ProSerAlaLeuProProValMet 88
    |||||
Db 296 CCTGTAGTCTCTCCCTCATCATG 319

RESULT 5
US-09-404-879A-5
; Sequence 5, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-404-879A-5

Alignment Scores:
Pred. No.: 1,46e-46 Length: 531
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
Gaps: 0
DB: 4

US-09-720-934-2_COPY_15_102 (1-88) x US-09-404-879A-5 (1-531)

QY 1 TTPAlaIleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
    |||||
Db 84 TGGGCTATTACATGTAAGAACGACTAAGCATGATTAACAAGTTGATTAACCTCAAACT 143
QY 21 ILeSerGlyPheIleThrGlyAspGlnAlaArgsnPhePheGlnSerGlyLeuPro 40
    |||||
Db 144 TCAGAGGTATACATAAAGGTGATCAAGCCCGTACTTTTCCACAGCTCAGGCTCTCCG 203
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetCsnAsnAspGlyArgMetAsp 60
    |||||
Db 204 GCCCGGTTTAGCTGAATAATGGGCTTATCAGATCTGAACAGAGATGGGAGATGTGAC 263
QY 61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
    |||||
Db 264 CAGCAAGAGTTCTTATAGCTATGAACTCATCAAGTTAAAGTTGCAAGGCCAACAGCTG 323
QY 81 ProSerAlaLeuProProValMet 88
    |||||
Db 324 CCTGTAGTCTCTCCCTCATCATG 347

RESULT 6
US-09-338-933-5
```

```
; Sequence 5, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; TITLE OF INVENTION: OVARIAN CANCER
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338,933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-5

Alignment Scores:
Pred. No.: 1,46e-46 Length: 531
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
Gaps: 0
DB: 4

US-09-720-934-2_COPY_15_102 (1-88) x US-09-338-933-5 (1-531)

QY 1 TTPAlaIleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
    |||||
Db 84 TGGGCTATTACATGTAAGAACGACTAAGCATGATTAACAAGTTGATTAACCTCAAACT 143
QY 21 ILeSerGlyPheIleThrGlyAspGlnAlaArgsnPhePheGlnSerGlyLeuPro 40
    |||||
Db 144 TCAGAGGTATACATAAAGGTGATCAAGCCCGTACTTTTCCACAGCTCAGGCTCTCCG 203
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetCsnAsnAspGlyArgMetAsp 60
    |||||
Db 204 GCCCGGTTTAGCTGAATAATGGGCTTATCAGATCTGAACAGAGATGGGAGATGTGAC 263
QY 61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
    |||||
Db 264 CAGCAAGAGTTCTTATAGCTATGAACTCATCAAGTTAAAGTTGCAAGGCCAACAGCTG 323
QY 81 ProSerAlaLeuProProValMet 88
    |||||
Db 324 CCTGTAGTCTCTCCCTCATCATG 347

RESULT 7
US-09-215-681-5
; Sequence 5, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; TITLE OF INVENTION: OF OVARIAN CANCER
; FILE REFERENCE: 210121.463
; CURRENT APPLICATION NUMBER: US/09/215,681A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-215-681-5

Alignment Scores:
Pred. No.: 1,46e-46 Length: 531
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
```

Best Local Similarity: 76.14%
Query Match: 78.21%
DB: 4

Mismatches: 13
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) * US-09-216-681-5 (1-531)

```
QY 1 TTPAlaIleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 84 TGGGCTATTACATCTGAAGAACTGACTAAGCATGATTAACAGTTTGAATCACTCAAACT 143
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 144 TCAGGAGGTTACATCAAGGTATCAAGCCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 203
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 204 GCCCGGTTTACGTAATATGCGCTTATCAGATCTGAACAGATGCGAAGATGAGAC 263
QY 61 GlnValGluPheSerIleAlaMetLysLeuLysLeuGlnGlyTyrGlnLeu 80
DB 264 CAGCAAGAGTTCTCTATGCTATGAAACTCATCAAGTTGACAGGCGCAACAGCTG 323
QY 81 ProSerAlaLeuProProValMet 88
DB 324 CCTGTAGTCTCTCTCTTATCATG 347
```

RESULT 8

US-09-216-003A-5
Sequence 5, Application US/09216003A
Patent No. 6670463
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Frudakis, Tony N.
APPLICANT: King, Gordon E.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
FILE REFERENCE: 210121.462
CURRENT APPLICATION NUMBER: US/09/216, 003A
CURRENT FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 310
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 531
TYPE: DNA
ORGANISM: Homo sapiens
US-09-216-003A-5

Alignment Scores:

Pred. No.: 1,466-46
Score: 359.00
Percent Similarity: 85.23%
Best Local Similarity: 76.14%
Query Match: 78.21%
DB: 4

Length: 531
Matches: 67
Conservative: 8
Mismatches: 13
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) * US-09-216-003A-5 (1-531)

```
QY 1 TTPAlaIleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 84 TGGGCTATTACATCTGAAGAACTGACTAAGCATGATTAACAGTTTGAATCACTCAAACT 143
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 144 TCAGGAGGTTACATCAAGGTATCAAGCCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 203
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 204 GCCCGGTTTACGTAATATGCGCTTATCAGATCTGAACAGATGCGAAGATGAGAC 263
QY 61 GlnValGluPheSerIleAlaMetLysLeuLysLeuGlnGlyTyrGlnLeu 80
DB 264 CAGCAAGAGTTCTCTATGCTATGAAACTCATCAAGTTGACAGGCGCAACAGCTG 323
QY 81 ProSerAlaLeuProProValMet 88
```

DB 324 CCTGTAGTCTCTCTCTTATCATG 347

RESULT 9
US-09-404-879A-72

Sequence 72, Application US/09404879A
Patent No. 6468546
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
APPLICANT: Algate, Paul A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.462C2
CURRENT APPLICATION NUMBER: US/09/404, 879A
CURRENT FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 393
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 72
LENGTH: 2017
TYPE: DNA
ORGANISM: Homo sapien
US-09-404-879A-72

Alignment Scores:
Pred. No.: 1,026-45
Score: 359.00
Percent Similarity: 85.23%
Best Local Similarity: 76.14%
Query Match: 78.21%
DB: 4

Length: 2017
Matches: 67
Conservative: 8
Mismatches: 13
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) * US-09-404-879A-72 (1-2017)

```
QY 1 TTPAlaIleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 84 TGGGCTATTACATCTGAAGAACTGACTAAGCATGATTAACAGTTTGAATCACTCAAACT 143
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 144 TCAGGAGGTTACATCAAGGTATCAAGCCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 203
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 204 GCCCGGTTTACGTAATATGCGCTTATCAGATCTGAACAGATGCGAAGATGAGAC 263
QY 61 GlnValGluPheSerIleAlaMetLysLeuLysLeuGlnGlyTyrGlnLeu 80
DB 264 CAGCAAGAGTTCTCTATGCTATGAAACTCATCAAGTTGACAGGCGCAACAGCTG 323
QY 81 ProSerAlaLeuProProValMet 88
DB 324 CCTGTAGTCTCTCTCTTATCATG 347
```

RESULT 10

US-09-338-933-72
Sequence 72, Application US/09338933
Patent No. 6488931
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer Lynn
APPLICANT: King, Gordon E.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
OVARIAN CANCER
FILE REFERENCE: 210121.462C1
CURRENT APPLICATION NUMBER: US/09/338, 933
CURRENT FILING DATE: 1999-06-23
NUMBER OF SEQ ID NOS: 312
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 72
LENGTH: 2017
TYPE: DNA
ORGANISM: Homo sapien
US-09-338-933-72

Alignment Scores:

Pred. No.: 1.02e-45 Length: 2017
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: 4 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-338-933-72 (1-2017)

QY 1 TTPAlaIeThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 84 TGGGCTATTACATCTGAAGAACGTAAGCATGATTAACAGTTGATTAACCTAAACCT 143
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 144 TCAGAGAGTTACATAACAGGATCAAGCCGCTACTTTTCCACAGTCAAGTCTGCCG 203
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 204 GCCCGGTTTATGCTGAATATGGCCCTTATCAGATCTGAACAAAGATGGGAAGATGAC 263
QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
DB 264 CAGCAAGAGTTCTCTATAGCTATGAACTCATCAAGTTAAAGTTGACAGGCCAACAGCTG 323
QY 81 ProSerAlaLeuProProValMet 88
DB 324 CCTGATGCTCCCTCCTATCATG 347

RESULT 11
US-09-215-681-72
Sequence 72, Application US/09215681A
Patent No. 6528253

GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Frudakis, Tony N.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
TITLE OF INVENTION: OF OVARIAN CANCER
FILE REFERENCE: 210121.463
CURRENT APPLICATION NUMBER: US/09/215.681A
CURRENT FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 310
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 72
LENGTH: 2017
TYPE: DNA
ORGANISM: Homo sapien
US-09-215-681-72

Alignment Scores:

Pred. No.: 1.02e-45 Length: 2017
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: 4 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-215-681-72 (1-2017)

QY 1 TTPAlaIeThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 84 TGGGCTATTACATCTGAAGAACGTAAGCATGATTAACAGTTGATTAACCTAAACCT 143
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 144 TCAGAGAGTTACATAACAGGATCAAGCCGCTACTTTTCCACAGTCAAGTCTGCCG 203
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 204 GCCCGGTTTATGCTGAATATGGCCCTTATCAGATCTGAACAAAGATGGGAAGATGAC 263

QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
DB 264 CAGCAAGAGTTCTCTATAGCTATGAACTCATCAAGTTAAAGTTGACAGGCCAACAGCTG 323
QY 81 ProSerAlaLeuProProValMet 88
DB 324 CCTGATGCTCCCTCCTATCATG 347

RESULT 12
US-09-216-003A-72
Sequence 72, Application US/09216003A
Patent No. 6670463

GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Frudakis, Tony N.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
FILE REFERENCE: 210121.462
CURRENT APPLICATION NUMBER: US/09/216.003A
CURRENT FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 310
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 72
LENGTH: 2017
TYPE: DNA
ORGANISM: Homo sapiens
US-09-216-003A-72

Alignment Scores:

Pred. No.: 1.02e-45 Length: 2017
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: 4 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-216-003A-72 (1-2017)

QY 1 TTPAlaIeThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 84 TGGGCTATTACATCTGAAGAACGTAAGCATGATTAACAGTTGATTAACCTAAACCT 143
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 144 TCAGAGAGTTACATAACAGGATCAAGCCGCTACTTTTCCACAGTCAAGTCTGCCG 203
QY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 264 CAGCAAGAGTTCTCTATAGCTATGAACTCATCAAGTTAAAGTTGACAGGCCAACAGCTG 323
QY 81 ProSerAlaLeuProProValMet 88
DB 324 CCTGATGCTCCCTCCTATCATG 347

RESULT 13
US-08-095-737-3

Sequence 3, Application US/08095737
Patent No. 5487979
GENERAL INFORMATION:
APPLICANT: DiFiore, Pier P
APPLICANT: Razioli, Francesca
TITLE OF INVENTION: A Substrate for the Epidermal Growth
TITLE OF INVENTION: Factor Receptor Kinase
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobb, Martens, Olson & Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach

```
STATE: California
COUNTRY: United States of America
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/095,737
FILING DATE: 19930722
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH060.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3033 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 111..2802
US-08-095-737-3

Alignment Scores:
Pred. No.: 1.75e-21 Length: 3033
Score: 205.00 Matches: 34
Percent Similarity: 68.18% Conservative: 26
Best Local Similarity: 38.64% Mismatches: 28
Query Match: 44.66% Indels: 0
DB: 1 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) * US-08-095-737-3 (1-3033)
QY 1 TTPAAlIeThVaGlUGlUaRgAlaYshSaSPGInGlnPhehIseRleuYsPro 20
Db 474 TGGGCTGTAAGTCTGAAGATTAAGCAATATGATGCAATTTTTCACGTTTAAGCCA 533
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
Db 534 GTGAGTGTGATTTTCTGCTGATTAAGCAACAGTGTGCTCACTCAAGTTACT 593
QY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaaspMetAsnAsnAspGlyArgMetAap 60
Db 594 GTGGAATCTCTTGGAAGATTGGAGATTGAGTGAATATGACCAAGATGAAAGCTGGAC 653
QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
Db 654 AGAGATGAGTTTGCAAGTTCGATGTTTTCGTATACGTGACAGTGAAGAAAGACCTGTG 713
QY 81 ProSerAlaLeuProProValMet 88
Db 714 CCAATGCTCTGCTCCACGCTTG 737

RESULT 14
US-08-480-145-3
/ Sequence 3, Application US/08480145
/ Patent No. 5717067
/ GENERAL INFORMATION:
/ APPLICANT: DiFiore, Pier P
/ APPLICANT: Fazio, Francesca
/ TITLE OF INVENTION: A Substrate for the Epidermal Growth
/ Factor Receptor Kinase
/ NUMBER OF SEQUENCES: 4
```

```
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive, Sixteenth floor
CITY: Newport Beach
STATE: California
COUNTRY: United States of America
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,145
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/095,737
FILING DATE: 22-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH060.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3033 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 111..2802
US-08-480-145-3

Alignment Scores:
Pred. No.: 1.75e-21 Length: 3033
Score: 205.00 Matches: 34
Percent Similarity: 68.18% Conservative: 26
Best Local Similarity: 38.64% Mismatches: 28
Query Match: 44.66% Indels: 0
DB: 1 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) * US-08-480-145-3 (1-3033)
QY 1 TTPAAlIeThVaGlUGlUaRgAlaYshSaSPGInGlnPhehIseRleuYsPro 20
Db 474 TGGGCTGTAAGTCTGAAGATTAAGCAATATGATGCAATTTTTCACGTTTAAGCCA 533
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
Db 534 GTGAGTGTGATTTTCTGCTGATTAAGCAACAGTGTGCTCACTCAAGTTACT 593
QY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaaspMetAsnAsnAspGlyArgMetAap 60
Db 594 GTGGAATCTCTTGGAAGATTGGAGATTGAGTGAATATGACCAAGATGAAAGCTGGAC 653
QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
Db 654 AGAGATGAGTTTGCAAGTTCGATGTTTTCGTATACGTGACAGTGAAGAAAGACCTGTG 713
QY 81 ProSerAlaLeuProProValMet 88
Db 714 CCAATGCTCTGCTCCACGCTTG 737

RESULT 15
US-08-477-389-3
/ Sequence 3, Application US/08477389
```

```

: Patent No. 5872219
: GENERAL INFORMATION:
: APPLICANT: DiFiore, Pier P
: APPLICANT: Fazioli, Francesca
: TITLE OF INVENTION: A Substrate for the Epidermal Growth
: FACTOR RECEPTOR KINASE
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Knobbe, Martens, Olson & Bear
: STREET: 620 Newport Center Drive, Sixteenth Floor
: CITY: Newport Beach
: STATE: California
: COUNTRY: United States of America
: ZIP: 92660
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/477,389
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/095,737
: FILING DATE: 22-JUL-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Israelsen, Ned A
: REGISTRATION NUMBER: 29,655
: REFERENCE/DOCKET NUMBER: NIH060.001A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 235-8550
: TELEFAX: (619) 235-0176
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3033 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 111..2802
: US-08-477-389-3

Alignment Scores:
Pred. No.: 1.75e-21 Length: 3033
Score: 205.00 Matches: 34
Percent Similarity: 68.18% Conservative: 26
Best Local Similarity: 38.64% Mismatches: 28
Query Match: 44.66% Indels: 0
DB: 2 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-08-477-389-3 (1-3033)

QY 1 TTPAlaIleThrValGluGluArgAlaIysHisAspGlnGlnPheHisSerLeuIysPro 20
Db 474 TGGGCTGTAAAGTCTGAAGATAAAGCCAAATATGATGCAATTTTGAACAGTTTAAGCCA 533

QY 21 ILeSerGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
Db 534 GTGGATGATATTTTGTCTGGGTGATTAAGTGAAACCACTGTTGCTCAACTTAAGTTACT 593

QY 41 GlnProValIleuAlaGlnIleTTPAlaIleuAlaAspMetAsnAspGlyArgMetAsp 60
Db 594 GTGGAATCTCTGGAAGAAGATTGGAGATTGATATATGACCAAGATGGAAGAAGCTGAC 653

QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
Db 654 AGAGATGATTTGACAGTTGACATGTTTGTATATCTGTGACATGAGAAAGAACTGTG 713

```

```

QY 81 ProSerAlaLeuProProValMet 88
Db 714 CCAATGTCCTTGCTCCAGCCTTG 737

Search completed: July 1, 2004, 19:53:37
Job time : 66.6206 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:09:20 ; Search time 269.977 Seconds
(without alignments)
1572.471 Million cell updates/sec

Title: US-09-720-934-2_COPY_15_102
Perfect score: 459
Sequence: 1 WAITBERAHHQDFHSLKP.....KLTKLQGYQLSPALPPVM 88

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3163042 segs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/p/US09720934/runat.30062004.064541.13481/app.query.fasta_1.1386
-DB=Published Applications NA -OPM=fastat -SUFFIX=mpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=bloms62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09720934 @CGN 1 1 1500 @runat.30062004.064541_13481
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT_DSPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -NARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubna/US10C_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
------------	-------	--------------------	----	-------------

1	459	100.0	3319	11	US-09-764-875-88	Sequence 88, Appl
2	459	100.0	3466	16	US-10-158-057-33	Sequence 33, Appl
3	425.5	92.7	568	10	US-09-764-881-55	Sequence 55, Appl
4	425.5	92.7	568	11	US-09-764-875-404	Sequence 404, Appl
5	425.5	92.7	568	13	US-09-764-881-55	Sequence 55, Appl
6	425.5	92.7	568	16	US-10-242-747-55	Sequence 55, Appl
7	425.5	92.7	568	16	US-10-158-057-127	Sequence 127, Appl
8	365	79.5	5828	13	US-10-398-8854-15	Sequence 15, Appl
9	359	78.2	480	9	US-09-884-441-60	Sequence 60, Appl
10	359	78.2	480	10	US-09-907-969-60	Sequence 60, Appl
11	359	78.2	480	10	US-09-827-271-60	Sequence 60, Appl
12	359	78.2	531	15	US-10-198-053-60	Sequence 60, Appl
13	359	78.2	531	9	US-09-884-441-55	Sequence 5, Appl
14	359	78.2	531	10	US-09-907-969-5	Sequence 5, Appl
15	359	78.2	531	10	US-09-827-271-5	Sequence 5, Appl
16	359	78.2	531	15	US-10-198-053-5	Sequence 5, Appl
17	359	78.2	2017	9	US-09-884-441-72	Sequence 72, Appl
18	359	78.2	2017	10	US-09-907-969-72	Sequence 72, Appl
19	359	78.2	2017	10	US-09-827-271-72	Sequence 72, Appl
20	359	78.2	2017	15	US-10-198-053-72	Sequence 72, Appl
21	351	76.5	462	10	US-09-918-995-21728	Sequence 21728, A
22	339	73.9	967	10	US-09-764-881-50	Sequence 50, Appl
23	339	73.9	967	11	US-09-764-875-411	Sequence 411, Appl
24	339	73.9	967	13	US-09-764-881-50	Sequence 50, Appl
25	339	73.9	967	16	US-10-242-747-50	Sequence 50, Appl
26	339	73.9	967	16	US-10-158-057-40	Sequence 40, Appl
27	212	46.2	292	9	US-09-864-761-20261	Sequence 20261, A
28	212	46.2	304	9	US-09-864-761-21373	Sequence 21373, A
29	212	46.2	310	9	US-09-864-761-19751	Sequence 19751, A
30	212	46.2	310	9	US-09-864-761-19759	Sequence 19759, A
31	212	46.2	466	9	US-09-864-761-3491	Sequence 3491, Ap
32	212	46.2	466	9	US-09-864-761-4631	Sequence 4631, Ap
33	199.5	43.5	605	12	US-10-152-119A-1126	Sequence 1126, Ap
34	199.5	43.5	860	13	US-10-027-632-165796	Sequence 165796, A
35	199.5	43.5	860	13	US-10-027-632-165797	Sequence 165797, A
36	199.5	43.5	860	16	US-10-027-632-165796	Sequence 165796, A
37	199.5	43.5	860	16	US-10-027-632-165797	Sequence 165797, A
38	199.5	43.5	3348	9	US-09-312-762A-2	Sequence 2, Appl
39	199.5	43.5	3485	15	US-10-171-581-354	Sequence 354, Appl
40	199.5	43.5	14707	9	US-09-312-762A-3	Sequence 3, Appl
41	198	43.1	662	13	US-10-027-632-290824	Sequence 290824, A
42	198	43.1	662	16	US-10-027-632-290824	Sequence 290824, A
43	197.5	43.0	3635	9	US-09-312-762A-6	Sequence 6, Appl
44	193.5	42.2	551	15	US-10-029-386-25205	Sequence 25205, A
45	190.5	41.5	3508	9	US-09-312-762A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-764-875-88
; Sequence 88, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 88
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-875-88

Alignment Scores: 1.41e-62 Length: 3319
Pred. No.: 459.00 Matches: 88
Score: 100.00% Conservative: 0

```
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 11
Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) * US-09-764-875-88 (1-3319)

QY 1 TTPAAlleThrValGluGlnArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 201 TGGGCATTAAGTCTAGAGAAAGAGGAGCATGATCAGAGTTCATGTTAAAGCCA 260

QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 261 ATATCTGATTCATTAATCTGCTGATCAAGCTAGAACTTTTTCATCTGGTTACCT 320

QY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 321 CAACCTGTTTACACAGATATGGGACATGCTGACATGATATATATGGAAGATGAT 380

QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeu 80
DB 381 CAAGTGAAGTTTCCATAGCTATGAAACTTATCAAACTGAGCTACAGGATATCAGCTA 440

QY 81 ProSerAlaLeuProProValMet 88
DB 441 CCTCTGCACTTCCCTCTGTCAATG 464

RESULT 2
US-10-158-057-33
; Sequence 33, Application US/10158057
; Publication No. US20040014039A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ205C1
; CURRENT APPLICATION NUMBER: US/10/158,057
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33
; LENGTH: 3466
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3194)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3465)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-158-057-33

Alignment Scores:
Pred. No.: 1,5e-62 Length: 3466
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) * US-10-158-057-33 (1-3466)

QY 1 TTPAAlleThrValGluGlnArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 319 TGGGCATTAAGTCTAGAGAAAGAGGAGCATGATCAGAGTTCATGTTAAAGCCA 378

QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 379 ATATCTGATTCATTAATCTGCTGATCAAGCTAGAACTTTTTCATCTGGTTACCT 438

QY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 441 CCTCTGCACTTCCCTCTGTCAATG 464
```

```
DB 439 CAACCTGTTTACACAGATATGGGACATGCTGACATGATATATGATGGAAGATGAT 498
QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeu 80
DB 499 CAAGTGAAGTTTCCATAGCTATGAAACTTATCAAACTGAGCTACAGGATATCAGCTA 558

QY 81 ProSerAlaLeuProProValMet 88
DB 559 CCTCTGCACTTCCCTCTGTCAATG 582

RESULT 3
US-09-764-881-55
; Sequence 55, Application US/09764881
; Publication No. US20030125246A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; PRIOR APPLICATION DATA REMOVED - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (536)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (556)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (562)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-881-55

Alignment Scores:
Pred. No.: 2,81e-58 Length: 568
Score: 425.50 Matches: 87
Percent Similarity: 98.86% Conservative: 0
Best Local Similarity: 98.86% Mismatches: 1
Query Match: 92.70% Indels: 1
DB: 10 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) * US-09-764-881-55 (1-568)

QY 1 TTPAAlleThrValGluGlnArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 120 TGGGCATTAAGTCTAGAGAAAGAGGAGCATGATCAGAGTTCATGTTAAAGCCA 177

QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 178 ATATCTGATTCATTAATCTGCTGATCAAGCTAGAACTTTTTCATCTGGTTACCT 237

QY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 238 CAACCTGTTTACACAGATATGGGACATGCTGACATGATATATGATGGAAGATGAT 297

QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeu 80
DB 298 CAAGTGAAGTTTCCATAGCTATGAAACTTATCAAACTGAGCTACAGGATATCAGCTA 357

QY 81 ProSerAlaLeuProProValMet 88
DB 358 CCTCTGCACTTCCCTCTGTCAATG 381

RESULT 4
```



```
US-09-764-875-404
; Sequence 404, Application US/09764875
; Publication No. US2004001893A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 404
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (536)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (556)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (562)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-875-404

Alignment Scores:
Pred. No.: 2,81e-58 Length: 568
Score: 425.50 Matches: 87
Percent Similarity: 98.86% Conservative: 0
Best Local Similarity: 98.86% Mismatches: 1
Query Match: 92.70% Indels: 1
DB: 11 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-764-875-404 (1-568)
QY 1 TTPALAIETHTVVALGUGLUARGALALYSHISASPGLNGINPHEHISERLEUYSPro 20
DB 120 TGGGCATTAACCTGTAGAGAGAAAGACGA--CATGATCGACGATTCATAGTTTAAAGCCA 177
QY 21 ILESERGLYPHEIETHTGILYASPGINAAARGASNPHEPHEGLINSEKGLYLEUPro 40
DB 178 ATATCTGATTCATTACTCGGTGATCAAGCTAGAACTTTTTCATCTGGGTTACT 237
QY 41 GLNPROVALLEUALAGINLIETTPALALEUALAASPMTASNASASGLYARGMETasp 60
DB 238 CAACCTGTTTTAGACAGATATGGGCCTAGCTGACATGAATATAGATGGAAGATGGAT 297
QY 61 GINVALGHPHESERILALAMETLYSEULIYLSLEULYSLGUNGILYTYrGInleu 80
DB 298 CAAGTGGAGTTTCCATAGCTATGAACTTATCAAACTGAACTCAAGATATACGCTA 357
QY 81 PROSERALALEUPROPROVALMET 88
DB 358 CCTCTGCACCTTCCCCCTGTGATG 381

RESULT 5
US-09-764-881-55
; Sequence 55, Application US/09764881
; Publication No. US20020086821A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (536)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (556)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (562)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-881-55

Alignment Scores:
Pred. No.: 2,81e-58 Length: 568
Score: 425.50 Matches: 87
Percent Similarity: 98.86% Conservative: 0
Best Local Similarity: 98.86% Mismatches: 1
Query Match: 92.70% Indels: 1
DB: 13 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-764-881-55 (1-568)
QY 1 TTPALAIETHTVVALGUGLUARGALALYSHISASPGLNGINPHEHISERLEUYSPro 20
DB 120 TGGGCATTAACCTGTAGAGAGAAAGACGA--CATGATCGACGATTCATAGTTTAAAGCCA 177
QY 21 ILESERGLYPHEIETHTGILYASPGINAAARGASNPHEPHEGLINSEKGLYLEUPro 40
DB 178 ATATCTGATTCATTACTCGGTGATCAAGCTAGAACTTTTTCATCTGGGTTACT 237
QY 41 GLNPROVALLEUALAGINLIETTPALALEUALAASPMTASNASASGLYARGMETasp 60
DB 238 CAACCTGTTTTAGACAGATATGGGCCTAGCTGACATGAATATAGATGGAAGATGGAT 297
QY 61 GINVALGHPHESERILALAMETLYSEULIYLSLEULYSLGUNGILYTYrGInleu 80
DB 298 CAAGTGGAGTTTCCATAGCTATGAACTTATCAAACTGAACTCAAGATATACGCTA 357
QY 81 PROSERALALEUPROPROVALMET 88
DB 358 CCTCTGCACCTTCCCCCTGTGATG 381

RESULT 6
US-10-242-747-55
; Sequence 55, Application US/10242747
; Publication No. US20040005577A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207C1
; CURRENT APPLICATION NUMBER: US/10/242,747
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,881
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
```

```

; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217 496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225 447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218 290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 55
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (536)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (556)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (562)
; OTHER INFORMATION: n equals a,t,g, or c
; US-10-242-747-55

```

Alignment Scores:

```

Pred. No.: 2,81e-58 Length: 568
Score: 425.50 Matches: 87
Percent Similarity: 98.868 Conservative: 0
Best Local Similarity: 98.868 Mismatches: 1
Query Match: 92.704 Indels: 1
DB: 16 Gaps: 0

```

US-09-720-934-2_COPY_15_102 (1-88) * US-10-242-747-55 (1-568)

```

QY 1 TTPAIAIEThVAlGluGluArgAlaYSHiSAsPGInGlnPheHISerLeuYsPro 20
DB 120 TGGGCGCATTAACGTAGAGAAAGAGCGA--CATGATCAGCAGTTCCATGTTTAAAGCCA 177
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
DB 178 ATATCTGATTAATCACTGATCAAGCTAGAAAATTITTTTTCATCTGGGTTACCT 237
QY 41 GlnProValLeuAlaGlnIleTTPAIAIEuAlaAspMetAsnAspGlyArgMetAsp 60
DB 238 CAACCTGTTTTCAGCAGATATGGGCACTAGCTGACATATAATGATGAAAGATGAT 297
QY 61 GlnValGlnPheSerIleAlaMetCysLeuIleCysLeuYsLeuGlnGlnGlyTyrgInLeu 80
DB 298 CAAGTGGAGTTTTCATACCTATGAACCTTATCAAACTGAACTGAAAGATATCAGCTA 357
QY 81 ProSerAlaLeuProProValMet 88
DB 358 CCTCTGACCTTCCCTGCTCATG 381

```

RESULT 7

```

US-10-158-057-127
; Sequence 127, Application US/10158057
; Publication No. US20040014039A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1205C1
; CURRENT APPLICATION NUMBER: US/10/158, 057
; PRIOR FILING DATE: 2002-06-12
; Prior Application removed - See File Wrapper or Palm

```

```

; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 127
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (536)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (556)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (562)
; OTHER INFORMATION: n equals a,t,g, or c
; US-10-158-057-127

```

Alignment Scores:

```

Pred. No.: 2,81e-58 Length: 568
Score: 425.50 Matches: 87
Percent Similarity: 98.868 Conservative: 0
Best Local Similarity: 98.868 Mismatches: 1
Query Match: 92.704 Indels: 1
DB: 16 Gaps: 0

```

US-09-720-934-2_COPY_15_102 (1-88) * US-10-158-057-127 (1-568)

```

QY 1 TTPAIAIEThVAlGluGluArgAlaYSHiSAsPGInGlnPheHISerLeuYsPro 20
DB 120 TGGGCGCATTAACGTAGAGAAAGAGCGA--CATGATCAGCAGTTCCATGTTTAAAGCCA 177
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro 40
DB 178 ATATCTGATTAATCACTGATCAAGCTAGAAAATTITTTTTCATCTGGGTTACCT 237
QY 41 GlnProValLeuAlaGlnIleTTPAIAIEuAlaAspMetAsnAspGlyArgMetAsp 60
DB 238 CAACCTGTTTTCAGCAGATATGGGCACTAGCTGACATATAATGATGAAAGATGAT 297
QY 61 GlnValGlnPheSerIleAlaMetCysLeuIleCysLeuYsLeuGlnGlnGlyTyrgInLeu 80
DB 298 CAAGTGGAGTTTTCATACCTATGAACCTTATCAAACTGAACTGAAAGATATCAGCTA 357
QY 81 ProSerAlaLeuProProValMet 88
DB 358 CCTCTGACCTTCCCTGCTCATG 381

```

RESULT 8

```

US-10-398-885A-15
; Sequence 15, Application US/10398885A
; Publication No. US20040053282A1
; GENERAL INFORMATION:
; APPLICANT: Sugita, Yuji
; APPLICANT: Hashida, Ryolchi
; APPLICANT: Ogawa, Kaoru
; APPLICANT: Nagasu, Takeshi
; APPLICANT: Obayashi, Masaya
; APPLICANT: Saito, Hirohisa
; APPLICANT: Takahashi, Eiki
; TITLE OF INVENTION: Method of Testing For Allergic Diseases
; FILE REFERENCE: SHIMIZU-07907
; CURRENT APPLICATION NUMBER: US/10/398, 885A
; PRIOR FILING DATE: 2003-08-11
; CURRENT APPLICATION NUMBER: PCT/JP01/08937
; PRIOR FILING DATE: 2001-10-11
; Prior Application removed - See File Wrapper or Palm

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Alignment Scores:	
Pred. No.:	3,89e-47
Score:	365.00
Percent Similarity:	86.3%
Best Local Similarity:	77.2%
Query Match:	79.5%
DB:	13
Length:	582
Matches:	68
Conservative:	8
Mismatches:	12
Indels:	0
Gaps:	0

US-09-720-934-2 COPY 15 102 (1-88) X US-10-398-885A-15 (1-5828)

QY	1	TPRALALEHRAVLGLUCINLRGALALAYSHLSAPSGINGINPHEHISERILEUYS	Pro 20
Db	85	TGGCTATTACCTCGTCAAGACGGTACTAGATGACAGACGATTGGATTAACCTAA	Act 144
QY	21	ILSERGLYPHEILETHRGLYASPGINALATGASNPHEPHEGLINSEGLYLEUP	Pro 40
Db	145	TCAGAGGTTACATTAACAGGTGATCAAGACCGTAATTTTTCCTCAATTCAGGTC	GCG 204
QY	41	GLAPROVALLEULAGINILETPRALALEULAAAPMEVASTASAPGLAIRMETASP	Met 60
Db	205	GCCCCCTTTTACCTGAATAATGGGCTTTTATACAGCTTAACACAGATGGGAAGT	GAT 264
QY	61	GLINVALGHPHESERILEALAWETIUSLEUILIETUSLEUYSLEUINGLIYTRGLIN	Leu 80
Db	265	CACCAAGAGTTCCTCCATTACCTATGAAACATCAACATGAAGCTTCAAGGCCAAC	GTTG 324
QY	81	PROSERALEUPROPROVALMET	Met 88
Db	325	CCGTGTGTTCTCCCTCTATTATG	348

```

RESULT 9
US-09-884-441-60
; Sequence 60, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Carter, Patrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.46CCT
; CURRENT APPLICATION NUMBER: US/09/884,441
; CURRENT FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-884-441-60

```

Alignment Scores:	
Pred. No.:	1.06e-47
Score:	359.00
Percent Similarity:	85.23%
Best Local Similarity:	76.14%
Query Match:	78.21%
DB:	9
Length:	480
Matches:	67
Conservative:	8
Mismatches:	13
Indels:	0
Gaps:	0

US-09-720-934-2_COPY_15_102 (1-88) X US-09-884-441-60 (1-480

[illegible]

```

RESULT 10
US-09-907-969-60
: Sequence 60, Application US/09907969
: Publication NO. US20030091580A1
: GENERAL INFORMATION:
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: King, Gordon F.
: APPLICANT: Algate, Paul A.
: APPLICANT: Fling, Steven P.
: APPLICANT: Retter, Marc W.
: APPLICANT: Fanger, Gary Richard
: APPLICANT: Reed, Steven G.
: APPLICANT: Vedvick, Thomas S.
: APPLICANT: Carter, Darrick
: APPLICANT: Hill, Paul
: APPLICANT: Albome, Earl
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
: FILE REFERENCE: 210121.46208
: CURRENT APPLICATION NUMBER: US/09/907,969
: CURRENT FILING DATE: 2001-07-17
: NUMBER OF SEQ. ID NOS: 596
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 60
: LENGTH: 480
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-907-969-60

```

Alignment Scores:	
Pred. No.:	1.06e-47
Score:	359.00
Percent Similarity:	85.23%
Best Local Similarity:	76.14%
Query Match:	78.21%
DB:	10
Length:	480
Matches:	67
Conservative:	8
Mismatches:	13
Indels:	0
Gaps:	0

US-09-720-934-2_COPY_15_102 (1-88) X US-09-907-969-60 (1-480)

Qy	TPRALAILThrvAlGIuLIuTrvAlIuYshIAspGInGInPPhnIISetLeuIuYsro	20
	
Db	TGGCTATTACATCATCGAAGACGTCATAGCATGATATAACAGTTTGATTAACCTCAACT	115
	
Qy	21 ILSErGlyPhehIePhcIuYsGInAlIaArGaAnPhePheGInSevGlyLeuPro	40
	
Db	116 TCGAGAGATCTACATACAGAGTATCAAGCCGCTACTTTTCTCTACAGTCAAGTGTGGC	175
	
Qy	41 GInProValIeuAlaGInIleTrrpAlaIeuAlaAspMeAsnaAsnAspGlyAlaMetAsp	60
	
Db	176 GCCCGCGATTTCCTGAATATGGCGCTTATCGATCTGGAACAAGATGGGAAGATGGAC	235

```

OY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeu 80
DB 236 CAGCAAGAGTTCTCTAATGCTATGAACTCATCAAGTTAAAGTTCAGGCGCAACAGCTG 295
OY 81 ProSerAlaLeuProProValMet 88
DB 296 CCTGTAGTCCTCCCTCATCATG 319

RESULT 11
US-09-827-271-60
; Sequence 60, Application US/09827271
; Publication No. US20030165504A1
; GENERAL INFORMATION:
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C6
; CURRENT APPLICATION NUMBER: US/09/827,271
; CURRENT FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 60
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-827-271-60

Alignment Scores:
Pred. No.: 1.06e-47 Length: 480
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-827-271-60 (1-480)

OY 1 TTPAlaIleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 56 TGGGCTATTACATCTGAAGAAGTCTAGCATGATGAACAGTTGATGTAACCTCAAACT 115
OY 21 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 116 TCAGAGGTTACATACAGAGTATCAAGCCCTACTTTTCTTACAGTCAGTCTGCCG 175
OY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 176 GCCCGGTTTATGCTGAATAATGGCCCTTATCAGATCTGAACAAGATGGGAAGATGGAC 235
OY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeu 80
DB 236 CAGCAAGAGTTCTCTAATGCTATGAACTCATCAAGTTAAAGTTCAGGCGCAACAGCTG 295
OY 81 ProSerAlaLeuProProValMet 88
DB 296 CCTGTAGTCCTCCCTCATCATG 319

RESULT 12
US-10-198-053-60
; Sequence 60, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198,053
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 624

```

```

; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-053-60

Alignment Scores:
Pred. No.: 1.06e-47 Length: 480
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: 15 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-10-198-053-60 (1-480)

OY 1 TTPAlaIleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 56 TGGGCTATTACATCTGAAGAAGTCTAGCATGATGAACAGTTGATGTAACCTCAAACT 115
OY 21 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
DB 116 TCAGAGGTTACATACAGAGTATCAAGCCCTACTTTTCTTACAGTCAGTCTGCCG 175
OY 41 GlnProValLeuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 176 GCCCGGTTTATGCTGAATAATGGCCCTTATCAGATCTGAACAAGATGGGAAGATGGAC 235
OY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuGlnGlyTyrGlnLeu 80
DB 236 CAGCAAGAGTTCTCTAATGCTATGAACTCATCAAGTTAAAGTTCAGGCGCAACAGCTG 295
OY 81 ProSerAlaLeuProProValMet 88
DB 296 CCTGTAGTCCTCCCTCATCATG 319

RESULT 13
US-09-884-441-5
; Sequence 5, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Carter, Derrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C7
; CURRENT APPLICATION NUMBER: US/09/884,441
; CURRENT FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-884-441-5

Alignment Scores:
Pred. No.: 1.22e-47 Length: 531
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-884-441-5 (1-531)

OY 1 TTPAlaIleThrValGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
DB 84 TGGGCTATTACATCTGAAGAAGTCTAGCATGATGAACAGTTGATGTAACCTCAAACT 143
OY 21 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40

```

```
Db 144 TCAGAGATTACATTAACAGGTGATCAAGCCGACTTTTCTTCTACAGTCAGGTCGCG 203
QY 41 GlnProValIleuAagInIleTTPAlaIleuAaAspMeKsRnsApGlyArqMetAsp 60
Db 204 GCCCGGTTTACTGAAATATGGCCCTTATCAGATCTGAACAAGATGGAGATGGAC 263
QY 61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
Db 264 CAGCAAGAGTTCTCTATAGCTATGAACTCATCAAGTTAAAGTTGACGGGCCAACAGCTG 323
QY 81 ProSerAlaIleuProProValMet 88
Db 324 CCTGTAGTCCTCCCTCCTATCATG 347

RESULT 14
US-09-907-969-5
; Sequence 5, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Rling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Hall, Paul
; APPLICANT: Albone, Earl
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C8
; CURRENT APPLICATION NUMBER: US/09/907.969
; NUMBER OF SEQ ID NOS: 536
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-907-969-5

Alignment Scores:
Pred. No.: 1.22e-47 Length: 531
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 13
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-907-969-5 (1-531)
QY 1 TTPAlaIleThrValGlnGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
Db 84 TGGGCTATTACATCTGAAGAACGACTAAAGCATATAACAGTTTGATTAACCTCAAACT 143
QY 21 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnDhePheGlnSerGlyLeuPro 40
Db 144 TCAGAGGTTTACATTAACAGGTGATCAAGCCGTTCTTTTCTTCAAGTCAGGTCGCG 203
QY 41 GlnProValIleuAagInIleTTPAlaIleuAaAspMeKsRnsApGlyArqMetAsp 60
Db 204 GCCCGGTTTACTGAAATATGGCCCTTATCAGATCTGAACAAGATGGAGATGGAC 263
QY 61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
Db 264 CAGCAAGAGTTCTCTATAGCTATGAACTCATCAAGTTAAAGTTGACGGGCCAACAGCTG 323
QY 81 ProSerAlaIleuProProValMet 88
Db 324 CCTGTAGTCCTCCCTCCTATCATG 347
```

```
RESULT 15
US-09-827-271-5
; Sequence 5, Application US/09827271
; Publication No. US20030165504A1
; GENERAL INFORMATION:
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C6
; CURRENT APPLICATION NUMBER: US/09/827.271
; CURRENT FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-827-271-5

Alignment Scores:
Pred. No.: 1.22e-47 Length: 531
Score: 359.00 Matches: 67
Percent Similarity: 85.23% Conservative: 8
Best Local Similarity: 76.14% Mismatches: 13
Query Match: 78.21% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x US-09-827-271-5 (1-531)
QY 1 TTPAlaIleThrValGlnGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
Db 84 TGGGCTATTACATCTGAAGAACGACTAAAGCATATAACAGTTTGATTAACCTCAAACT 143
QY 21 IleSerGlyPheIleThrGlyAspGlnAlaArgAsnDhePheGlnSerGlyLeuPro 40
Db 144 TCAGAGGTTTACATTAACAGGTGATCAAGCCGTTCTTTTCTTCAAGTCAGGTCGCG 203
QY 41 GlnProValIleuAagInIleTTPAlaIleuAaAspMeKsRnsApGlyArqMetAsp 60
Db 204 GCCCGGTTTACTGAAATATGGCCCTTATCAGATCTGAACAAGATGGAGATGGAC 263
QY 61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
Db 264 CAGCAAGAGTTCTCTATAGCTATGAACTCATCAAGTTAAAGTTGACGGGCCAACAGCTG 323
QY 81 ProSerAlaIleuProProValMet 88
Db 324 CCTGTAGTCCTCCCTCCTATCATG 347
```

Search completed: July 1, 2004, 20:15:43
Job time : 275.977 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 11:41:07 ; Search time 271.625 Seconds

(without alignments)
1376.313 Million cell updates/sec

Title: US-09-720-934-2_COPY_15_102

Perfect score: 459

Sequence: 1 WAITVERAHDQGFHSLKP.....KLTKLQGYQLPSALPPVM 88

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 3373863 seqs, 212409941 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xip
-Q=/cgm2_1/USPRO.spool_p/US097070934/runat_30062004_064539_13422/epg_query.fasta_1.1386
-DB=N_Geneseq_29Jan04 -OPMT=fastcap -SUFFIX=rng -MINMATCH=0.1 -LOOPCT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosome2 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US09720934_0CGN_1_1_1596_0runat_30062004_064539_13422 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOBFRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002s:*\n7: geneseqn2003as:*\n8: geneseqn2003bs:*\n9: geneseqn2003cs:*\n10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	459	100.0	877	4	AAK93179 Human CDN
2	459	100.0	877	4	AAK91610 Human CDN
3	459	100.0	1676	4	AAK94611 Human SH3
4	459	100.0	2079	2	AAZ34573 Human SH3
5	459	100.0	2131	4	AAK94139 Human Full
6	459	100.0	2131	4	AAH16578 Human CDN
7	459	100.0	3319	4	ABK43498 DNA encod
8	459	100.0	3466	4	AAI63825 Human pol

9	459	100.0	5195	2	AAZ34572 Human SH3
10	459	100.0	5199	2	AAZ34570 Human SH3
11	459	100.0	5458	2	AAZ34571 Human SH3
12	453	98.7	7435	5	AAK84763 DNA encod
13	441	96.1	3723	3	AAZ39009 Mouse ESE
14	441	96.1	5082	3	AAZ39008 Mouse ESE
15	441	96.1	5144	3	AAZ39025 Mouse ESE
16	441	96.1	5738	3	AAZ39024 Mouse ESE
17	426	92.8	676	4	AAH08146 Human CDN
18	425.5	92.7	568	4	AAI63919 Human pol
19	425.5	92.7	568	4	AAI63919 Human pol
20	425.5	92.7	568	4	AAI63919 Human pol
21	365	79.5	5828	6	AAI47247 Allergic
22	365	79.5	6103	4	AAK52332 Human pol
23	359	78.2	480	3	AAA69750 Human ova
24	359	78.2	480	6	ABN72644 Ovarian c
25	359	78.2	480	8	ADA08809 Human ova
26	359	78.2	531	3	AAA69695 Human ova
27	359	78.2	531	6	ABN72589 Ovarian c
28	359	78.2	531	8	ADA08754 Human ova
29	359	78.2	2017	3	AAA69762 Human ova
30	359	78.2	2017	6	ABN72656 Ovarian c
31	359	78.2	2017	8	ADA08821 Human ova
32	359	78.2	3593	3	AAZ39011 Mouse ESE
33	359	78.2	4625	3	AAZ39010 Mouse ESE
34	359	78.2	4975	3	AAZ39027 Mouse ESE
35	359	78.2	6014	3	AAZ39026 Mouse ESE
36	351	76.5	462	8	ACH34516 Human end
37	339	73.9	967	4	AAI63832 Human pol
38	339	73.9	967	4	AAI63832 Human pol
39	339	73.9	967	4	ABK43821 DNA encod
40	291	63.4	4447	4	AAK02055 DNA encod
41	262	57.1	3981	4	ABL01995 Drosophila
42	240	52.3	7225	4	ABL01994 Drosophila
43	212	46.2	292	4	AAI22800 Probe #12
44	212	46.2	292	4	ABK43807 Human fo
45	212	46.2	292	4	AAI48103 Probe #16

ALIGNMENTS

RESULT 1
AAK93179
ID AAK93179 standard; CDN: 877 BP.
AC AAK93179;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human cDNA clone representative sequence, SEQ ID NO: 1639.
XX
KW Human, full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
EN EPI130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
XX
PR 11-JAN-2000; 2000JP-00118774.
XX
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use

PT in genetic manipulation.

XX Example 11: SEQ ID NO 1639; 1380bp + Sequence Listing; English.

CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5' and 3' ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesized by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence was used as the representative sequence
 CC from a human clone which was used in homology searches to identify the
 CC clone. Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in CD-ROM format directly from
 CC EPO

CC Sequence 877 BP; 208 A; 228 C; 223 G; 213 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.:	2,446-58	Length:	877
Score:	459.00	Matches:	88
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-720-934-2_COPY_15_102 (1-88) x AAK93179 (1-877)

QY 1 TTPAIAIETHThVaIGluGluArgAlaYSHiASpGInGlnPheHiseSerLeuYsPro 20
 DB 247 TGGGCGATTAACGTAGAGAAAGAGGAGCATGATCAGCATTCATGTTAAAGCCA 306
 QY 21 ILESerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
 DB 307 ATATCTGATTCATTACTGATGATCAAGCTAGAACTTTTTCATCTGGTTACT 366
 QY 41 GlnProValLeuAlaGlnIleTTPAIAleuAlaAspMetAsnAsnAspGlyArgMetLeu 60
 DB 367 CAACCTGTTTTCAGACAGATATGGGCACTGATGATCAATTAATGATGAAAGATGGAT 426
 QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
 DB 427 CAAGTGGAGTTTTCATAGCTATGAAACTTATCAACTGAACTCAAGATATACGCTA 486
 QY 81 ProSerAlaLeuProProValMet 88
 DB 487 CCCTCTGCACCTCCCTGTCATG 510

RESULT 2
 AAK91610
 ID AAK91610 standard; cDNA; 877 BP.

XX AAK91610;

DT 06-NOV-2001 (first entry)

DE Human cDNA 5'-end sequence, SEQ ID NO: 70.

KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

OS Homo sapiens.

PN EPI130094-A2.

PD 05-SEP-2001.

PF 07-JUL-2000; 2000EP-00114089.

PR 08-JUL-1999; 99JP-00194486.

PR 11-JAN-2000; 2000JP-00118774.

PR 02-MAY-2000; 2000JP-00183765.

XX (HELI-) HELIX RES INST.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI: 2001-52425/58.

PT 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.

PS Claim 2; SEQ ID NO 70; 1380bp + Sequence Listing; English.

CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5' and 3' ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesized by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is the nucleotide sequence of the 5'-end of
 CC a cDNA provided in the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in CD-
 CC ROM format directly from EPO

CC Sequence 877 BP; 208 A; 228 C; 223 G; 213 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.:	2,446-58	Length:	877
Score:	459.00	Matches:	88
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-720-934-2_COPY_15_102 (1-88) x AAK91610 (1-877)

QY 1 TTPAIAIETHThVaIGluGluArgAlaYSHiASpGInGlnPheHiseSerLeuYsPro 20
 DB 247 TGGGCGATTAACGTAGAGAAAGAGGAGCATGATCAGCATTCATGTTAAAGCCA 306
 QY 21 ILESerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
 DB 307 ATATCTGATTCATTACTGATGATCAAGCTAGAACTTTTTCATCTGGTTACT 366
 QY 41 GlnProValLeuAlaGlnIleTTPAIAleuAlaAspMetAsnAsnAspGlyArgMetLeu 60
 DB 367 CAACCTGTTTTCAGACAGATATGGGCACTGATGATCAATTAATGATGAAAGATGGAT 426
 QY 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
 DB 427 CAAGTGGAGTTTTCATAGCTATGAAACTTATCAACTGAACTCAAGATATACGCTA 486
 QY 81 ProSerAlaLeuProProValMet 88
 DB 487 CCCTCTGCACCTCCCTGTCATG 510

RESULT 3
 AAK94611
 ID AAK94611 standard; cDNA; 1676 BP.

XX AAK94611;

DT 07-NOV-2001 (first entry)

DE Human cDNA, SEQ ID NO: 3565.

KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

OS Homo sapiens.

PN EPI130094-A2.


```
XX 05-SEP-2001.
PD
XX
XX 07-JUL-2000; 2000EP-00114089.
PF
XX 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX WPI; 2001-524255/58.
DR
XX P-PSDB; AAM93676.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
in genetic manipulation.
XX
XX Disclosure; SEQ ID NO 3565; 1380bp + Sequence listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
XX clones. 830 cDNA molecules encoding a human protein have been isolated
XX and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX been determined. Primers for synthesizing the full length cDNA are useful
XX for clarifying the function of the protein encoded by the cDNA. The full
XX length clones were obtained by construction of full length enriched cDNA
XX libraries that were synthesised by the oligo-capping method. The primers
XX enable the production of the full length cDNA easily without any special
XX methods. The present sequence is a human cDNA provided in the
XX specification. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in CD-ROM format directly
XX from EPO
XX
XX Sequence 1676 BP; 485 A; 381 C; 458 G; 352 T; 0 U; 0 Other;
SQ
XX
XX Alignment Scores:
XX Pred. No.: 5.91e-58 Length: 1676
XX Score: 459.00 Matches: 88
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 4 Gaps: 0
XX
XX US-09-720-934-2_COPY_15_102 (1-88) x AA234611 (1-1676)
XX
XX 1 TnpalalrthValGlugluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
Db 306 TGGGCGCATTAAGTGAAGAAAGCCAGCAATGATGCGAGTTCATGTTTAAAGCCA 365
XX
XX 21 TleserGlyPheIlethnrglyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
Db 366 ATATCTGGAATTCATTACTGCGATGATCAAGCTAGAAACCTTTTTTTCATCATCGGTTACT 425
XX
XX 41 GlnproValleuAlaGlnIleTnpAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
Db 426 CCACTCTGTTTGAACAAGATATGGGCACTACCTACATGAATATGATGAAGATGAT 485
XX
XX 61 GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
Db 486 CAACTGGAATTCATTCATGATGATGAATTTCAAACTGAAGCTACAAAGATATCAGCTA 545
XX
XX 81 ProSerAlaLeuProProValMet 88
Db 546 CCTCTGCACTTCCCTGTCATG 569
XX
XX
XX RESULT 4
XX ID AA234573 standard; cDNA; 2079 BP.
XX AC AA234573;
XX
```

```
DT 01-FEB-2000 (first entry)
XX
XX Human SH3D1A cDNA clone 5.
DE
XX
XX SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
KW megakaryocytic abnormality; myeloproliferative disorder;
KW platelet disorder; neural disorder; thrombocytopenia;
KW haematopoietic disorder; cognitive dysfunction; macrocephaly;
KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 136..2079
FT CDS /*tag= a
XX
XX WC9953062-A2.
XX
XX 21-OCT-1999.
XX
XX 16-APR-1999; 99WO-US008371.
XX
XX 16-APR-1998; 98US-0082007P.
XX
XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.
XX
XX Korenberg JR, Chen X;
XX
XX WPI; 1999-633829/54.
XX
XX P-PSDB; AAY32157.
XX
XX Nucleic acid from the human SH3D1A gene and its products, useful for the
XX diagnosis and treatment of myeloproliferative disorders and leukemia.
XX
XX Claim 2; Fig 12; 99JP; English.
XX
XX This is the nucleotide sequence of a cDNA clone, termed clone 5,
XX corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
XX contributes to the development of platelets and the pathogenesis of
XX leukaemias, both in general and in particular those involving the
XX megakaryocytic lineage. The SH3D1A gene maps to the small candidate
XX region for low platelets on chromosome 21. Sequencing of 5 different
XX sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
XX least 3 isoforms exist. The invention provides methods for the diagnosis
XX and treatment of megakaryocytic abnormality, myeloproliferative disorder,
XX platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
XX association of gains in chromosome 21 with leukaemias, neural
XX abnormalities, dysfunctions and disorders including brain malformations
XX and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
XX and colpocephaly. Methods are also provided for: suppressing cells unable
XX to regulate themselves; screening for a somatic alteration in the SH3D1A
XX gene; monitoring the progress and adequacy of a treatment; monitoring
XX tumour risk progress or megakaryocytic abnormality, myeloproliferative
XX disorder, haematopoietic disorder, platelet disorder or leukaemia; and
XX treatment of a subject (including a prenatal subject) having
XX megakaryocytic abnormality, myeloproliferative disorder, platelet
XX disorder, leukaemia or neural disorder using a nucleic acid that
XX expresses SH3D1A or its antisense nucleic acid
XX
XX Sequence 2079 BP; 700 A; 451 C; 495 G; 433 T; 0 U; 0 Other;
SQ
XX
XX Alignment Scores:
XX Pred. No.: 7.93e-58 Length: 2079
XX Score: 459.00 Matches: 88
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 2 Gaps: 0
XX
XX US-09-720-934-2_COPY_15_102 (1-88) x AA234573 (1-2079)
XX
XX 1 TnpalalrthValGlugluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro 20
XX
```

```

Db      178 TGGGCCATACGTGAGAGAAAGAGGAGCATGATCAGACATTCATGTTAAAGCCA 237
Qy      21 11eSerGlyPhe11eThrGlyaspGln1aAargAsnPhenPheGlnSerGlyLeuPro 40
Db      238 ATATCTGAGTTTCATCTGATGATCAAGCTAAGAACTTTTTCATCTGGGTTACCT 297
Qy      41 GlnProValLeuAlaGln11eTrrPa1aLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
Db      298 CAACCTGTTTTCAGCAGATATGGGCACTAGCTGACATGATATGATGAGAAATGAT 357
Qy      61 GlnValGluPheSer11eAlaMetLysLeu11eLysLeuGlnGlyTyrGlnLeu 80
Db      358 CAAGTGGAGTTTTCATGACTATGAAGCTTATCAAGCTAAGAACTATACAGCTA 417
Qy      81 ProSerAlaLeuProProValMet 88
Db      418 CCTCTGCACTTCCCCCTGTCTCATG 441

```

RESULT 5

AAK94139 standard; cDNA; 2131 BP.

AAK94139;
06-NOV-2001 (first entry)

Human full-length cDNA, SEQ ID NO: 2646.

Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

Homo sapiens.

EP1130094-A2.

05-SEP-2001.

07-JUL-2000; 2000BP-00114089.

08-JUL-1999; 99JP-00194486.

11-JAN-2000; 2000JP-0018774.

02-MAY-2000; 2000JP-00183765.

(HELI-) HELIX RES INST.

Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

WPI; 2001-524255/58.

P-PsDB; AAM93229.

830 Primers useful for synthesizing full length cDNA clones and their use

in genetic manipulation.

Claim 8; SEQ ID NO 2646; 1380bp + Sequence Listing; English.

The invention relates to primers for synthesizing full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5' and 3'-ends of the cDNA molecules have been determined. Primers for synthesizing the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a full length human cDNA of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly from EPO

Sequence 2131 BP; 694 A; 469 C; 526 G; 442 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8.2e-58

Length: 2131

Score: 459.00 Matches: 88
Percent Similarity: 100.00%
Percent Identity: 100.00%
Query Match: 100.00%
DB: 4
Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x AAK94139 (1-2131)

```

Qy      1 TrrPa1a11eThrValGluGluArgAlaLysH1sAspGlnGlnPheHisSerLeuysPro 20
Db      247 TGGGCCATACGTGAGAGAAAGAGGAGCATGATCAGACATTCATGTTAAAGCCA 306
Qy      21 11eSerGlyPhe11eThrGlyaspGln1aAargAsnPhenPheGlnSerGlyLeuPro 40
Db      307 ATATCTGAGTTTCATCTGATGATCAAGCTAAGAACTTTTTCATCTGGGTTACCT 366
Qy      41 GlnProValLeuAlaGln11eTrrPa1aLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
Db      367 CAACCTGTTTTCAGCAGATATGGGCACTAGCTGACATGATATGATGAGAAATGAT 426
Qy      61 GlnValGluPheSer11eAlaMetLysLeu11eLysLeuGlnGlyTyrGlnLeu 80
Db      427 CAAGTGGAGTTTTCATGACTATGAAGCTTATCAAGCTAAGAACTATACAGCTA 486
Qy      81 ProSerAlaLeuProProValMet 88
Db      487 CCTCTGCACTTCCCCCTGTCTCATG 510

```

RESULT 6

AAH16578 standard; cDNA; 2131 BP.

AAH16578;

26-JUN-2001 (first entry)

Human cDNA sequence SEQ ID NO:15658.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000BP-00116126.

29-JUL-1999; 99JP-00248036.

27-AUG-1999; 99JP-00300253.

11-JAN-2000; 2000JP-0018776.

02-MAY-2000; 2000JP-00183765.

09-JUN-2000; 2000JP-00241899.

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J,
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

Claim 8; SEQ ID NO 15658; 2537bp + Sequence Listing; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AA892446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

CC
XX
SQ Sequence 2131 BP; 691 A; 448 C; 572 G; 420 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8.2e-58 Length: 2131
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-720-934-2_COPY_15_102 (1-88) x AAH16578 (1-2131)

QY 1 ThrAlaIethrValIGuGluArgAlaIySHISAspGInGInPheHisSerLeuLysPro 20
Db TGGGCCCAATACGTAGAGAGAAAGCCAGCATATACGACGTTCATGTTAAACCA 448
QY 21 ILeSerGIyPheIleThrGIyASpGInAlaArgAsnPhPheGInSerGIyLeuPro 40
Db ATATCTGGATTCTACTGCGATCAAGCTAGAACTTTTTCATCTGGGTTTACT 508
QY 41 GInProValIeuaIaGInIleTYPAlaIeuaIaAspMetAsnaAsnAspGIyArgMetAsp 60
Db CAACCTGTTTATAGACAGATATGGCGACTGCTGACATGTAATATGATGGAATGAT 568
QY 61 GlnValGlnPheSerIleAlaMetLysLeuIleLysLeuLysLeuGInGIyTyrGlnLeu 80
Db CAAGTGSAGTTTTCATAGCTATGAACCTTATCAACATGAAGATACAGCTA 628
QY 81 ProSerAlaIeuProProValMet 88
Db CCCCTGCACTCCCTCCGTGATG 652

RESULT 7
ABK43498
ID ABK43498 standard; cDNA; 3319 BP.
XX
XX ABK43498;
AC
XX
DT 05-JUN-2002 (first entry)
XX
DE DNA encoding novel central nervous system protein #78.
XX
KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy; gene; ss.
XX
XX Homo sapiens.
XX

PN W0200155318-A2.
XX
XX 02-AUG-2001.
PD
XX
PF 17-JAN-2001; 2001WO-US001332.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234937P.
PR 25-SEP-2000; 2000US-0234938P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR

PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236357P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 02-OCT-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251864P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI

XX MPI: 2001-581633/65.
DR P-PSDB; AAU87168.
DR
XX
PT New isolated nucleic acid encoding a protein for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives.
PT
XX
PS Claim 1; SEQ ID NO 88; 837pp; English.
XX
CC The invention describes an isolated nucleic acid molecule (I) encoding a
CC novel central nervous system protein. (I) and polypeptides (II) encoded
CC by (I), are used to treat a medical conditions and in diagnosis of a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC adenocarcinomas and irritable bowel syndrome, reproductive system
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infection. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,
Alignment Scores:
Pred. No.: 1.5e-57 Length: 3319
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-720-934-2_COPY_15_102 (1-88) x ABK43498 (1-3319)
QY 1 TTPAlaIleThrValGluGluArgAlaIleYSHISAspGlnInPheHISerLeuIysPro 20
Db 201 TGGGCACTAACTGTAGAGAAAGACGAGCATGATCGACAGTTCCATATGTTAAGCCA 260
QY 21 IleserGlyPheIleThrGlyAspGlnAlaArgAenPhePheGlnIserGlyLeuPro 40
Db 261 ATATCTGATTCATTAATCGTGTATCAAGCTAAGAACTTTTTCATCTCGGTTACT 320
QY 41 GlnProValIleuAlaGlnIleThrPalaleuAlaApmetAenAspGlyArgMetAap 60
Db 321 CAACCTGTTTGAACACATATGAGCAGTACGTCAGTAAGAAATGAGGAATGAT 380
QY 61 GlnValGluPheSerIleAlaMetIysLeuIleIysLeuIysLeuGlnIleYrGlnLeu 80
Db 381 CAAGTCAGATTTTCCATGACTATGAATCTTATCAAACTGAAGACTACAGAGATATCAGCTA 440
QY 81 ProSerAlaLeuProProValMet 88
Db 441 CCTCTGACACTTCCCCGTGTCATG 464
RESULT 8
AAI63825
ID AAI63825 standard; cDNA, 3466 BP.
XX
AC AAI63825;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 33.

PR	08-DEC-2000;	2000US-0251856P-	
PR	08-DEC-2000;	2000US-0251864P-	
PR	08-DEC-2000;	2000US-0251869P-	
PR	08-DEC-2000;	2000US-0251989P-	
PR	08-DEC-2000;	2000US-0251990P-	
PR	11-DEC-2000;	2000US-0254097P-	
PR	05-JAN-2001;	2001US-0259678P-	
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
XX			
PI	Rosen CA, Barash SC, Ruben SM,		
XX	WPI, 2001-488781/53.		
DR	P-FSDb; AAM43519.		
XX			
PT	New isolated nucleic acids and polypeptides, useful for diagnosing,		
PT	treating and/or preventing human diseases and disorders.		
XX			
PS	Claim 1; SEQ ID NO 33; 664pp +	Sequence Listing; English.	
XX			
CC	The invention relates to human polynucleotides (AA163803-AA164012) and the encoded proteins (AAM43497-AA43660) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes were isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as atherosclerotic ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published_pct_sequences		
XX	Sequence 3466 BP; 1056 A; 829 G; 861 G; 710 T; 0 U; 10 Other;		
SO			
Alignment Scores:			
	Pred. No.:	1.6e-57	Length: 3466
	Score:	459.00	Matches: 88
	Percent Similarity:	100.00%	Conservative: 0
	Best Local Similarity:	100.00%	Mismatches: 0
	Query Match:	100.00%	Indels: 0
	DB:	4	Gaps: 0
US-09-720-934-2_COPY_15_102 (1-88) x AA163825 (1-3466)			
OY	1	TTPALAIIETHRVAIGLUNARGALALYSHIASPGLINPHEHISERTLEUYS	Pro 20
Db	319	TGGGCATACACGTAGAGGAAAGAGCAATGATCAGACAGTTCATAGTTAAAGCCA	378
OY	21	IIIESERGLYPHEIIETHRGVYASPGNIAAYGASRPHPEPHEGINSERGLYLAU	Pro 40
Db	379	ATATCTGCATTCATTACTCTGCTGATCAAGCTAGAAATTTTTTTTCAATGTGGGTTA	CCT 438
OY	41	GLNPROVALLEUALAGINIIETTPALALEUALAASPMECAASNAASPGIAYRGME	ASP 60
Db	439	CAACCTGTTTTAGCAACAGATATCGGCACATAGCTGACATGATATATATGGA	AAATGGAT 498
OY	61	GINVALGIVPHSERITLALMELYSLEUUIELYSLEULYSLEUNGINGLYTYRGIN	LEU 80
Db	499	CAAGTGGAGTTTCCATAGCTATGAAACTTATCAAACTGAACTCAAGATATACGCTA	558
OY	81	PROSERIALALEUPROPROVALMET	88
Db	559	CCCTGTGCACCTTCCCCCTGTCTANG	582
RESULT 9			
AA234572			
AA234572 standard; cDNA; 5195 BP.			

XX	AA234572;		
AC			
XX	01-FEB-2000	(first entry)	
DT			
XX			
XX			
DE	Human SH3D1A cDNA clone 11.		
XX			
XX			
KM	SH3D1A gene; human; Down's syndrome; leukaemia; cancer;		
KM	megakaryocytic abnormality; myeloproliferative disorder;		
KM	platelet disorder; neural disorder; thrombocytopenia;		
KM	haematopoietic disorder; cognitive dysfunction; microcephaly;		
KM	lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.		
OS			
XX	Homo sapiens.		
PH			
PH	Key	Location/Qualifiers	
FT	CDS	239..3886	
FT		/*tag= a	
PN	W09953062-A2.		
PD	21-OCT-1999.		
PF	16-APR-1999;	99WO-US008371.	
PR	16-APR-1999;	98US-0082007P.	
XX			
PA	(CEDA-) CEDARS SINAI HEALTH SYSTEM.		
PI	Korenberg JR, Chen X;		
XX			
DR	WPI; 1999-633829/54.		
XX	P-PSDB; AA332156.		
PT	Nucleic acid from the human SH3D1A gene and its products, useful for the		
PT	diagnosis and treatment of myeloproliferative disorders and leukemia.		
XX			
PS	Claim 2; Fig 10; 99pp; English.		
XX			
CC	This is the nucleotide sequence of full-length cDNA (clone 11)		
CC	corresponding to a novel human SH3 gene, termed the SH3D1A gene, that		
CC	contributes to the development of platelets and the pathogenesis of		
CC	leukaemias, both in general and in particular those involving the		
CC	megakaryocytic lineage. The SH3D1A gene maps to the small candidate		
CC	region for low platelets on chromosome 21. Sequencing of 5 different		
CC	sizes of cDNA clone from fetal brain (see AA234570-74) suggests that at		
CC	least 3 isoforms exist. The invention provides methods for the diagnosis		
CC	and treatment of megakaryocytic abnormality, myeloproliferative disorder,		
CC	platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,		
CC	platelet disorder on chromosome 21, low platelets in deletion for 21,		
CC	association of gains in chromosome 21 with leukaemias, neural		
CC	abnormalities, dysfunctions and disorders including brain malformations		
CC	and corresponding cognitive dysfunctions, microcephaly, lissencephaly,		
CC	and colpocephaly. Methods are also provided for: suppressing cells unable		
CC	to regulate themselves; screening for a somatic alteration in the SH3D1A		
CC	gene; monitoring the progress and adequacy of a treatment; monitoring		
CC	tumour risk progress or megakaryocytic abnormality, myeloproliferative		
CC	disorder, haematopoietic disorder, platelet disorder or leukaemia; and		
CC	treatment of a subject (including a prenatal subject) having		
CC	megakaryocytic abnormality, myeloproliferative disorder, platelet		
CC	disorder, leukaemia or neural disorder using a nucleic acid that		
CC	expresses SH3D1A or its antisense nucleic acid		
XX			
SEQ	Sequence 5195 BP; 1592 A; 1148 C; 1237 G; 1218 T; 0 U; 0 Other;		
Alignment Scores:			
Pred. No.:	2.77e-57	Length:	5195
Score:	459.00	Matches:	88
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

```
US-09-720-934-2_COPY_15_102 (1-88) x AAZ34572 (1-5195)
QY 1 TTPAlallethThrValGluGluAAlaYSHisAspGlnGlnPheHisSerLeuYsPro 20
DB 281 TGGGCCATTAACCTAGAGGAAAGCGAACGATCATGATCAAGTTCATGTTTAAAGCCA 340
QY 21 TlleserGlyPheIleThrGlyAspGlnAlaArgAsnPhethePheGlnSerGlyLeuPro 40
DB 341 ATATCTGGATTCACTACTGGTGATCAAGCTAAACTTTTTCATCTGGGTTAACTCT 400
QY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 401 CAACCTGTTTATAGACAGATATGGGCACTAGCTACATGATATATGATGAAGATGAT 460
QY 61 GlnValGlnPheSerIleAlaMetYsLeuIleYsLeuYsLeuGlnGlyTyrGlnLeu 80
DB 461 CAAGTGGAGTTTCCATAGCTATGAACTTATCAAACTGAAGCTACAAGATATCAAGCTA 520
QY 81 ProSerAlaLeuProProValMet 88
DB 521 CCTCTGCACCTCCCTGTCATG 544
RESULT 10
AAZ34570
ID AAZ34570 standard; cDNA; 5199 BP.
AC AAZ34570;
DT 01-FEB-2000 (first entry)
XX
DE Human SH3D1A cDNA.
XX
KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
KW megakaryocytic abnormality; myeloproliferative disorder;
KW platelet disorder; neural disorder; thrombocytopenia;
KW hematopoietic disorder; cognitive dysfunction; microcephaly;
KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 208..3642
FT /*tag= a
XX
PN WO9953062-A2.
XX
PD 21-OCT-1999.
XX
PF 16-APR-1999; 99WC-US008371.
XX
PR 16-APR-1998; 98US-0082007P.
XX
PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
XX
PI Korenberg JR, Chen X;
XX
DR WPI, 1999-633829/54.
XX
DR P-PSDB; AAY32154.
XX
PT Nucleic acid from the human SH3D1A gene and its products, useful for the
PT diagnosis and treatment of myeloproliferative disorders and leukemia.
XX
PS Claim 2, Fig 5, 9pp; English.
XX
CC This is the nucleotide sequence of full-length cDNA corresponding to a
CC novel human SH3 gene, termed the SH3D1A gene, that contributes to the
CC development of platelets and the pathogenesis of leukemias, both in
CC general and in particular those involving the megakaryocytic lineage. The
CC SH3D1A gene maps to the small candidate region for low platelets on
CC chromosome 21. Sequencing of 5 different sizes of cDNA clones (see
CC AAZ34570-74) suggests that at least 3 isoforms exist. The invention
CC provides methods for the diagnosis and treatment of megakaryocytic
CC abnormality, myeloproliferative disorder, platelet disorder, acute
```

```
CC leukemia, neural disorders, thrombocytopenia, platelet disorder on
CC chromosome 21, low platelets in deletion for 21, association of gains in
CC disorders including brain malformations and corresponding cognitive
CC dysfunctions, microcephaly, lissencephaly, and colpocephaly. Methods are
CC also provided for: suppressing cells unable to regulate themselves;
CC screening for a somatic alteration in the SH3D1A gene; monitoring the
CC progress and adequacy of a treatment; monitoring tumor risk progress or
CC megakaryocytic abnormality, myeloproliferative disorder, hematopoietic
CC disorder, platelet disorder or leukemia; and treatment of a subject
CC (including a prenatal subject) having megakaryocytic abnormality,
CC myeloproliferative disorder, platelet disorder, leukaemia or neural
CC disorder using a nucleic acid that expresses SH3D1A or its antisense
CC nucleic acid
XX
SQ Sequence 5199 BP; 1615 A; 1152 C; 1214 G; 1218 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 2.78e-57 Length: 5199
Score: 459.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-720-934-2_COPY_15_102 (1-88) x AAZ34570 (1-5199)
QY 1 TTPAlallethThrValGluGluAAlaYSHisAspGlnGlnPheHisSerLeuYsPro 20
DB 250 TGGGCCATTAACCTAGAGGAAAGCGAACGATCATGATCAAGTTCATGTTTAAAGCCA 309
QY 21 TlleserGlyPheIleThrGlyAspGlnAlaArgAsnPhethePheGlnSerGlyLeuPro 40
DB 310 ATATCTGGATTCACTACTGGTGATCAAGCTAAACTTTTTCATCTGGGTTAACTCT 369
QY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaAspMetAsnAsnAspGlyArgMetAsp 60
DB 370 CAACCTGTTTATAGACAGATATGGGCACTAGCTACATGATATATGATGAAGATGAT 429
QY 61 GlnValGlnPheSerIleAlaMetYsLeuIleYsLeuYsLeuGlnGlyTyrGlnLeu 80
DB 430 CAAGTGGAGTTTCCATAGCTATGAACTTATCAAACTGAAGCTACAAGATATCAAGCTA 489
QY 81 ProSerAlaLeuProProValMet 88
DB 490 CCTCTGCACCTCCCTGTCATG 513
RESULT 11
AAZ34571
ID AAZ34571 standard; cDNA; 5458 BP.
AC AAZ34571;
XX
DT 01-FEB-2000 (first entry)
XX
DE Human SH3D1A cDNA clone 21.
XX
KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
KW megakaryocytic abnormality; myeloproliferative disorder;
KW platelet disorder; neural disorder; thrombocytopenia;
KW hematopoietic disorder; cognitive dysfunction; microcephaly;
KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 267..3929
FT /*tag= a
XX
PN WO9953062-A2.
XX
PD 21-OCT-1999.
XX
```



```
Db 305 TGGGCCATACTGTAGAGGAAGGAGCAGCATGATCAGCAGCTCCATGTTTAAAGCCA 364
QY 21 TLESerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
Db 365 ATATCTGATTCATCTACTGGTGGATCAAGCTAAGAACTTTTTCATCTGGGTTACT 424
QY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaAspMetAsnAspGlyArgMetAsp 60
Db 425 CAACCTGTTTATGACAGATATGGGCACTGACATGATATATGATGAGAGATGAT 484
QY 61 GlnValGlnPheSerIleAlaMetIleLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
Db 485 CAAGTGAAGTTTTCATAGCTATGATAAATCTTATCAAACTGAAGCTACAAAGATATCAGCTA 544
QY 81 ProSerAlaLeuProProValMet 88
Db 545 CCCTCTGCACTTCCCTGTCATG 568

RESULT 13
AAZ39009
ID AAZ39009 standard; cDNA; 3723 BP.
XX
AC AAZ39009;
XX
DT 28-FEB-2000 (first entry)
XX
DE Mouse Esei coding sequence.
XX
KM Mouse; murine; Esei; Esee2; endocytosis; vesicular trafficking;
KM EH-domain; actin cytoskeleton; detection; cancer; infection;
KM antiproliferative; antiviral; ss.
XX
OS Mus sp.
XX
PN WO9955728-A2.
XX
PD 04-NOV-1999.
XX
PF 27-APR-1999; 99WO-CA000375.
XX
PR 27-APR-1998; 98CA-00230201.
PR 05-FEB-1999; 99US-0118739P.
XX
PA (HSCR-) HSC RES & DEV LP.
XX
PI Egan SE, Wang W, Sengar A;
XX
DR WPI: 2000-052802/04.
DR P-PSDB; AAY57444.
XX
PT New nucleic acid encoding Esei and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.
XX
PS Claim 6; Page 40-42; 99pp; English.
XX
CC The present sequence encodes mouse Esei. The present invention
CC specifically describes mammalian Esei and 2 proteins (I) and their splice
CC variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (I)
CC are involved in regulation of clathrin-mediated endocytosis (as a complex
CC with Esp15 protein), vesicular trafficking and actin cytoskeleton.
CC Generally (I) (or its (ant)agonists, mimetics, fragments and inactive
CC mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
CC polynucleotide; agents that downregulate expression of Ese genes or
CC antagonists of an Ese binding partner are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of Esei is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (I) is used to promote endocytosis of selected cells. (Ant)agonists of
CC (I) or Ab are used to suppress abnormal proliferation of cells that can
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
```

```
CC cultures, by forming an Ese-Esp15 complex, then binding dynamn to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission
XX
SQ Sequence 3723 BP; 1065 A; 941 C; 1025 G; 692 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 8.52e-55 Length: 3723
Score: 441.00 Matches: 85
Percent Similarity: 97.73% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 2
Query Match: 96.08% Indels: 0
DB: 3 Gaps: 0
US-09-720-934-2_copy_15_102 (1-88) x AAZ39009 (1-3723)
QY 1 TTPAlaIleThrValGluGluArgAlaLysAspGlnGlnPheHisSerIleuLeuPro 20
Db 43 TGGCCATACTGTGAGGAAGGCCAAGCATGACCAAGATTCCTTAGCTGAGCCG 102
QY 21 TLESerGlyPheIleThrGlyAspGlnAlaArgAsnPhenPheGlnSerGlyLeuPro 40
Db 103 ATAGCGGATTTATTTACTGGTGANCAAGCAGAACTTTTTCATCTGGGTTACT 162
QY 41 GlnProValIleuAlaGlnIleTTPAlaLeuAlaAspMetAsnAspGlyArgMetAsp 60
Db 163 CAGCTGCTCTTAGACAAATATGGCGCTGCGCATGATAATACATGAGAGATGAT 222
QY 61 GlnValGlnPheSerIleAlaMetIleLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu 80
Db 223 CAAGTGAAGTTTTCATAGCCATGAAAGCTTATCAAACTGAAGCTACAAAGATATCAGCTC 282
QY 81 ProSerAlaLeuProProValMet 88
Db 283 CCCTCGCACTTCCCTGTCATG 306

RESULT 14
AAZ39008
ID AAZ39008 standard; cDNA; 5082 BP.
XX
AC AAZ39008;
XX
DT 28-FEB-2000 (first entry)
XX
DE Mouse Esei full length cDNA sequence.
XX
KM Mouse; murine; Esei; Esee2; endocytosis; vesicular trafficking;
KM EH-domain; actin cytoskeleton; detection; cancer; infection;
KM antiproliferative; antiviral; ss.
XX
OS Mus sp.
XX
PN WO9955728-A2.
XX
PD 04-NOV-1999.
XX
PF 27-APR-1999; 99WO-CA000375.
XX
PR 27-APR-1998; 98CA-00230201.
PR 05-FEB-1999; 99US-0118739P.
XX
PA (HSCR-) HSC RES & DEV LP.
XX
PI Egan SE, Wang W, Sengar A;
XX
DR WPI: 2000-052802/04.
DR P-PSDB; AAY57444.
XX
PT New nucleic acid encoding Esei and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.
XX
```

PS Claim 6; Page 38-40; 99pp; English.

XX The present sequence encodes mouse Ese1. The present invention
CC specifically describes mammalian Ese1 and 2 proteins (I) and their splice
CC variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (I)
CC are involved in regulation of clathrin-mediated endocytosis (as a complex
CC with Esp1 protein), vesicular trafficking and actin cytoskeleton.
CC Generally (I) (or its (ant)agonists, mimetics, fragments and inactive
CC mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
CC polynucleotide; agents that downregulate expression of Ese genes or
CC antagonists of an Ese binding partner are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of Ese1 is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (I) is used to promote endocytosis of selected cells. (Ant)agonists of
CC (I) or Ab are used to suppress abnormal proliferation of cells that can
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-Esp1 complex, then binding dynam to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signaling, tissue development or synaptic transmission
CC XX

Sequence 5082 BP; 1363 A; 1253 C; 1391 G; 1075 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,3e-54	Length:	5082
Score:	441.00	Matches:	85
Percent Similarity:	97.73%	Conservative:	1
Best Local Similarity:	96.59%	Mismatches:	2
Query Match:	96.08%	Indels:	0
DB:	3	Gaps:	0

US-09-720-934-2_COPY_15_102 (1-88) x AAZ39008 (1-5082)

QY	1	TrypAlaIeTnVAlGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro	20
DB	301	TGGGCGATTAACGTGGAGGAAAGGCCCAAGCATGACGACATTCCTTAAGCTGAAGCG	360
QY	21	IlleSerGlyPheIleTThGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro	40
DB	361	ATAGCGGATTTATTAATCGTGATCAAGCGAAGAACTTTTTCCAATCTGGTTACCT	420
QY	41	GlnProValIleuAlaGlnIleTTrpAlaIleuAlaAspMetAsnAsnArgGlyArgMetLasp	60
DB	421	CAGCTGTCTTAACCAATATAGGCGCTAGCGGACATGATTAACGATGGAAGATGAGAT	480
QY	61	GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu	80
DB	481	CAAGTGAATTTTCATAGCCATGAAGCTTATCAAACTGAAGCTTCAAGGATATCAGCTC	540
QY	81	ProSerAlaLeuProProValMet	88
DB	541	CCCTCCACACTCCCTCCCTGTCATG	564

RESULT 15

AAZ39025

ID AAZ39025 standard; cDNA, 5144 BP.

AAZ39025;

28-FEB-2000 (first entry)

Mouse Ese1 coding sequence.

Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;

EH-domain and SH3-domain regulator of endocytosis; cancer; infection;

antiproliferative; antiviral; ss.

Mus sp.

PN MO9955728-A2.

PD 04-NOV-1999.

PF 27-APR-1999; 99WO-CA000375.

PR 27-APR-1999; 98CA-02230201.

PS 05-FEB-1999; 99US-0118739P.

PA (HSCR-) HSC RES & DEV LP.

PI Egan SE, Wang W, Sengar A;

DR MPI: 2000-052802/04.

DR P-PSDB: AA57443.

New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of endocytosis, used e.g. for treating cancer or preventing viral infection.

Claim 6; Page 59-62; 99pp; English.

XX The present invention specifically describes mammalian Ese1 and 2
CC proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
CC regulator of endocytosis). (I) are involved in regulation of clathrin-
CC mediated endocytosis (as a complex with Esp1 protein), vesicular
CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
CC mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
CC sequences antisense to the (I) polynucleotide; agents that downregulate
CC expression of Ese genes or antagonists of an Ese binding partner are used
CC to treat diseases associated with undesirable endocytosis and resulting
CC changes in cellular function. Particularly overexpression of Ese1 is used
CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
CC administration of (I) is used to promote endocytosis of selected cells.
CC (Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
CC cells that can be stimulated to proliferate by a growth factor receptor;
CC and similar compounds (also inactive Ese mutants) can be used to prevent
CC viral infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-Esp1 complex, then binding dynam to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signaling, tissue development or synaptic transmission. The
CC present sequence represents mouse Ese1 coding sequence
CC XX

Sequence 5144 BP; 1447 A; 1298 C; 1415 G; 984 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,33e-54	Length:	5144
Score:	441.00	Matches:	85
Percent Similarity:	97.73%	Conservative:	1
Best Local Similarity:	96.59%	Mismatches:	2
Query Match:	96.08%	Indels:	0
DB:	3	Gaps:	0

US-09-720-934-2_COPY_15_102 (1-88) x AAZ39025 (1-5144)

QY	1	TrypAlaIeTnVAlGluGluArgAlaLysHisAspGlnGlnPheHisSerLeuLysPro	20
DB	43	TGGGCGATTAACGTGGAGGAAAGGCCCAAGCATGACGACATTCCTTAAGCTGAAGCG	102
QY	21	IlleSerGlyPheIleTThGlyAspGlnAlaArgAsnPhePheGlnSerGlyLeuPro	40
DB	103	ATAGCGGATTTATTAATCGTGATCAAGCGAAGAACTTTTTCCAATCTGGTTACCT	162
QY	41	GlnProValIleuAlaGlnIleTTrpAlaIleuAlaAspMetAsnAsnArgGlyArgMetLasp	60
DB	163	CAGCTGTCTTAACCAATATAGGCGCTAGCGGACATGATTAACGATGGAAGATGAGAT	222
QY	61	GlnValGluPheSerIleAlaMetLysLeuIleLysLeuLysLeuGlnGlyTyrGlnLeu	80
DB	223	CAAGTGAATTTTCATAGCCATGAAGCTTATCAAACTGAAGCTTCAAGGATATCAGCTC	282
QY	81	ProSerAlaLeuProProValMet	88

, Fri Jul 2 10:23:36 2004

us-09-720-934-2_copy_15_102.rng

Page 13

Db 283 CCTTCACACTCCCCCTGATG 306

Search completed: July 1, 2004, 13:31:27
Job time : 278.625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 12:16:50 ; Search time 2189.79 Seconds
(without alignments)
1900.151 Million cell updates/sec

Title: US-09-720-934-2_COPY_215_310
Perfect score: 501
Sequence: 1 WAPQSSRLKVRQLFNSHDK.....SGQPLPVLPPEYIPSPFR 96

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-O=/cg2.1/USPTO.spool.p/US09720934/runat.30062004.064539.13430/app.query.fasta_1.1386
-DB=GenEmbl -QFMT=fastap -SUPFIX=rge -MIMATCH=0.1 -LOOPT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NOR=excl -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720934.@CGN.1.1.7509.@runat.30062004.064539.13430 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-FEV_TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pin:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hngo_hum:*
40: em_hngo_mus:*
41: em_hngo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	501	100.0	1299	9 BC039036	BC039036 Homo sapi
2	501	100.0	1676	6 BD127640	BD127640 Primer fo
3	501	100.0	1676	9 AK075290	AK075290 Homo sapi
4	501	100.0	2079	6 BD205036	BD205036 Isolated
5	501	100.0	2131	6 AX880753	AX880753 Sequence
6	501	100.0	2131	6 BD127168	BD127168 Primer fo
7	501	100.0	2131	6 BD158570	BD158570 Primer fo
8	501	100.0	2131	9 AK027846	AK027846 Homo sapi
9	501	100.0	2131	9 AK074554	AK074554 Homo sapi
10	501	100.0	2199	9 BC058925	BC058925 Homo sapi
11	501	100.0	3723	10 AF132478	AF132478 Mus muscu
12	501	100.0	3812	10 AF132672	AF132672 Rattus no
13	501	100.0	4025	10 AF127798	AF127798 Rattus no
14	501	100.0	4321	9 HSM806384	HSM806384 Homo sapi
15	501	100.0	5145	10 AF132481	AF132481 Mus muscu
16	501	100.0	5195	6 BD205035	BD205035 Isolated
17	501	100.0	5199	6 BD205033	BD205033 Isolated
18	501	100.0	5287	9 AF064243	AF064243 Homo sapi
19	501	100.0	5381	9 AF114488	AF114488 Homo sapi
20	501	100.0	5458	6 BD205034	BD205034 Isolated
21	501	100.0	6439	9 AF114487	AF114487 Homo sapi
22	501	100.0	7247	9 AF064244	AF064244 Homo sapi
23	474	94.6	7105	5 AF032118	AF032118 Xenopus l
24	359.5	71.8	136937	2 AC139627	AC139627 Takifugu
25	349	69.7	2017	6 BD264687	BD264687 Compositi
26	349	69.7	2017	6 AR238090	AR238090 Sequence
27	349	69.7	2017	6 AR257631	AR257631 Sequence
28	349	69.7	2017	6 AR283677	AR283677 Sequence
29	349	69.7	2017	6 AX366305	AX366305 Sequence
30	349	69.7	3594	10 AF132480	AF132480 Mus muscu
31	349	69.7	4977	10 AF132479	AF132479 Mus muscu
32	349	69.7	5804	10 AK122480	AK122480 Mus muscu
33	346	69.1	1749	9 BC020921	BC020921 Homo sapi
34	346	69.1	2085	9 BC038963	BC038963 Homo sapi
35	345	68.9	4447	6 AX428899	AX428899 Sequence
36	345	68.9	4557	9 AF182199	AF182199 Homo sapi
37	345	68.9	5828	6 BD167848	BD167848 Method fo
38	345	68.9	5828	9 AF248540	AF248540 Homo sapi
39	345	68.9	5938	9 AB033082	AB033082 Homo sapi
40	345	68.9	6091	9 AF182198	AF182198 Homo sapi
41	292	58.3	3750	3 AF054612	AF054612 Drosophi
42	292	58.3	3999	3 AF053957	AF053957 Drosophi
43	292	58.3	127035	3 AC006574	AC006574 Drosophi
44	292	58.3	160282	3 AC009253	AC009253 Drosophi
45	292	58.3	273995	3 AF003669	AF003669 Drosophi

RESULT 1

ALIGNMENTS

BC039036 1299 bp mRNA linear PRI 24-DEC-2002
LOCUS BC039036
DEFINITION Homo sapiens, similar to intersectin 1 (SH3 domain protein), clone IMAGE:5784342, mRNA, partial cds.
ACCESSION BC039036
VERSION BC039036.1 GI:24659505
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1299)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Aklter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, R., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, D.C., McQuell, J., Pearson, R., Stancirpop, S., Thomas, P.D., Touchman, D.W., Tsurgren, C., Vogt, J.L., Walker, M.A., Weherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.liml.gov>
Series: IPAK Plate: 82 Row: 1 Column: 2
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4504796.
FEATURES
source
1..1299
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5784342"
/ciseu_type="uterus, leiomyosarcoma"
/clone_lib="NTH MGC_71"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
264..>1299
/codon_start=1
/product="similar to intersectin 1 (SH3 domain protein)"
/protein_id="AMH39036.1"
/db_xref="gi:24711335"
/translation="MAQPTTGGSLDIWATTVERRAKHDOFFSLKISGFTTDDQ
RNFPGSLPQPVLAQIWMALMNDGMDQVEFIANKLTKLQGSQLSALPVM
KOQPVASAPAFGAGISAMPPLTVAIVPMGSI.PVVGMSPTLVSVPTAIVPLAN
GAPVIGLPAPAFHAKATLKPSSSFRSGSGLNTKLQKASFPVASVPPVAEAVP
OSSRLKYQLFNSHKTMSGHLTGPOARTIMOSSLPOALASIMNLSDIDODGLTA
EETILAMTLDVAMSGCLPVLPPPIPPRRRVRSSGSISSIVSDQRLPPEPV
LEDEQOQKKKK"
ORIGIN
Alignment Scores:
Pred. No.: 1,21e-50 Length: 1299
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

DB: 9 Gaps: 0
US-09-720-934-2_COPY_215_310 (1-96) x BC039036 (1-1299)
QY 1 TTPALAVAPROGINSErSerArgLeuysTRYArgInLeuPheAsnSerHisAspLys 20
DB 906 TGGGCTGTTCCCTCACTCATCAAGACTGAAAATACAGGCAATTATTTCAATGATGACAGAAA 965
QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetClnSerSerLeu 40
DB 966 ACTATGAGTGTGACACTTCAACAGTCCCAAGCAAGCAACATCTTATGACAGTCAAGTTTA 1025
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
DB 1026 CCACAGGCTCAGCTGCTGCTTCAATATGAAATCTTTCTGACATTGATCAGATGGAACCTT 1085
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 1086 ACAGCAGAGGAATTATATCTGAGCAATGACACCTCATTTGATGTATGCTGCGCAACCA 1145
QY 81 LeuProProValLeuProProGlnTyrIleProProSerPheArgArg 96
DB 1146 CTGGCACCCTGCTCTGCTCCACAGAAATACATTCACCTCTTTAAGAGA 1193
RESULT 2
BD127640 1676 bp DNA linear PAT 18-SEP-2002
LOCUS BD127640
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127640
VERSION BD127640.1 GI:23222585
KEYWORDS JP 2002017375-A/3071.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1676)
Oca, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y., Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.
Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 3071 22-JAN-2002;
COMMENT HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/3071
PD 22-JAN-2002
PF 07-JUN-2000 JP 2000253172
PI TOSHIO OCA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, PC C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N5/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES
source
FT CDS Location/Qualifiers
1..1676
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1.6e-50 Length: 1676
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x BD127640 (1-1676)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTYrArgGlnLeuPheAsnSerHisAspLys 20
DB 906 TGGGCTGTTCTCAGTCATCAAGACCTGAAATACAGGCAATATTCATATGACAA 965

QY 21 ThMetSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
DB 966 ACTATGAGTGGACACTTAACAGGCTCCCAAGCAAGAACTATCTTATGACGTCAAGTTA 1025

QY 41 ProGlnAlaGlnLeuAlaSerIleTPAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
DB 1026 CCACAGGCTCAGCTGGCTTCAATATGAACTTTCTTGACATTCATCAAGATGAAAACTT 1085

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 1086 ACAGCAGAGGAATTTATCTGGCAATGACCTCATTTAGTATGATGTCGGCCAACTTA 1145

QY 81 LeuProProValLeuProProGlyTyrIleProProSerPheArgArg 96
DB 1146 CTGCCACCTGCTCCTGCCCTCCAGAAATACATTCCACTTTTGAAGA 1193

RESULT 3
AK075290 1676 bp mRNA linear PRI 03-SEP-2002
LOCUS AK075290
DEFINITION Homo sapiens cDNA FLJ30809 fis, clone Y79PA1000778, highly similar
to Homo sapiens intersecin long form mRNA.
ACCESSION AK075290.1 GI:22761283
VERSION AK075290.1
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS 1
Isegai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T., Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S., Kawai-Hio,Y., Saico,K., Yamamoto,J., Makamatsu,A., Nakamura,Y., Kojima,S., Nagahata,K., Masuno,Y., Ono,T., Okano,K., Yoshikawa,Y., Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and Ninomiya,K.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1676)
AUTHORS Isegai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (25-MAR-2002) Takao Isegai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@nri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo; Laboratory of Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).

FEATURES
source
1..1676
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Y79PA1000778"
/cell_line="Y79"
/cell_type="retinoblastoma"
/clone_lib="Y79PA1"
/note="cloning vector: pME18SFL3"

ORIGIN
Alignment Scores:
Pred. No.: 1.6e-50 Length: 1676
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AK075290 (1-1676)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTYrArgGlnLeuPheAsnSerHisAspLys 20
DB 906 TGGGCTGTTCTCAGTCATCAAGACCTGAAATACAGGCAATATTCATATGACAA 965

QY 21 ThMetSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
DB 966 ACTATGAGTGGACACTTAACAGGCTCCCAAGCAAGAACTATCTTATGACGTCAAGTTA 1025

QY 41 ProGlnAlaGlnLeuAlaSerIleTPAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
DB 1026 CCACAGGCTCAGCTGGCTTCAATATGAACTTTCTTGACATTCATCAAGATGAAAACTT 1085

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 1086 ACAGCAGAGGAATTTATCTGGCAATGACCTCATTTAGTATGATGTCGGCCAACTTA 1145

QY 81 LeuProProValLeuProProGlyTyrIleProProSerPheArgArg 96
DB 1146 CTGCCACCTGCTCCTGCCCTCCAGAAATACATTCCACTTTTGAAGA 1193

RESULT 4
BD205036 2079 bp DNA linear PAT 17-JUN-2003
LOCUS BD205036
DEFINITION Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.
ACCESSION BD205036.1 GI:33014806
VERSION BD205036.1
KEYWORDS JP 2002511267-A/4.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 2079)
Korenberg,J.R. and Chen,X.N.
TITLE Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof
JOURNAL Patent: JP 2002511267-A 4 16-APR-2002;
CEDARS SINAI HEALTH SYSTEM ET AL

COMMENT OS Homo sapiens (human)
PN JP 2002511267-A/4
PD 16-APR-2002
PR 16-APR-1999 JP 2000543610
PR 16-APR-1998 US 60/082007
PI JULIE R KORENBERG, XIAO NING CHEN
PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,
PC C12Q1/68,
PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and leukemia

FEATURES
source
1..2079
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 2.02e-50 Length: 2079
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) X BD205036 (1-2079)

QY 1 TtpalavalProglInSerSerArgLeuLysTyArgInLeuPheAsnSerHisAspLys 20
DB 778 TGGGCTGTTCTCCTCACTCATCAAGACTGAAATACAGCAATATTTCAATAGCATACAA 837
QY 21 ThrmetserglyHisLeuThrGlyProglInlaArgThrIleLeuMetGlnSerSerIeu 40
DB 838 ACTATGATGACACTTAACAGGTCCCAAGCAACACTATTTCTTATGACAGTCAAGTTTA 897
QY 41 ProglInlaGlnLeuAlaSerIleThrAsnLeuSerAspIleAspGlnAspGlyLysIeu 60
DB 898 CCACAGGCTCAGCTGCTTCAATATGAAATCTTCTGACATGATCAAGATGAAACTT 957
QY 61 ThrIaGluGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 958 ACAGAGAGGAATTTATCTGCAATGACACTCATGTATGATGTATGTCTGGCCAACTA 1017
QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArgArg 96
DB 1018 CTGCCACTGTCTGCTCCTCCAGAAATACATTCACCTCTTTAGAGAA 1065

RESULT 5
LOCUS AX880753 2131 bp DNA linear PAT 17-DEC-2003
DEFINITION Sequence 15658 from Patent EP1074617.
ACCESSION AX880753
VERSION AX880753.1 GI:40035489
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.,
Primers for synthesizing full-length cDNA and their use
JOURNAL Patent: JP 1074617-A 15658 07-FEB-2001;
Research Association for Biotechnology (JP)

FEATURES
source location/Qualifiers
1..2131
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
347..>2131
/note="unnamed protein product"

CDS
/codon_start=1
/protein_id="CAB92347.1"
/db_xref="GI:40035490"

ORIGIN
Alignment Scores:
Pred. No.: 2,07e-50 Length: 2131
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) X AX880753 (1-2131)

QY 1 TtpalavalProglInSerSerArgLeuLysTyArgInLeuPheAsnSerHisAspLys 20
DB 878 TGGGCTGTTCTCCTCACTCATCAAGACTGAAATACAGCAATATTTCAATAGCATACAA 937
QY 21 ThrmetserglyHisLeuThrGlyProglInlaArgThrIleLeuMetGlnSerSerIeu 40
DB 938 ACTATGATGACACTTAACAGGTCCCAAGCAACACTATTTCTTATGACAGTCAAGTTTA 997
QY 41 ProglInlaGlnLeuAlaSerIleThrAsnLeuSerAspIleAspGlnAspGlyLysIeu 60
DB 998 CCACAGGCTCAGCTGCTTCAATATGAAATCTTCTGACATGATCAAGATGAAACTT 1057
QY 61 ThrIaGluGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 1058 ACAGAGAGGAATTTATCTGCAATGACACTCATGTATGATGTATGTCTGGCCAACTA 1117
QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArgArg 96
DB 1118 CTGCCACTGTCTGCTCCTCCAGAAATACATTCACCTCTTTAGAGAA 1165

RESULT 6
LOCUS BD127168 2131 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127168
KEYWORDS BD127168.1 GI:23222113
JP 2002017375-A/2599.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2131)
Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
Primers for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 2599 22-JAN-2002;
HELIOS RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/2599

COMMENT
PD 22-JAN-2002 JP 2000253172
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/PC
10,
PC C12P21/02, C12O1/66//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key

FEATURES
source location/Qualifiers
1..2131
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 2,07e-50 Length: 2131
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) X BD127168 (1-2131)

QY 1 TtpalavalProglInSerSerArgLeuLysTyArgInLeuPheAsnSerHisAspLys 20
|||||

Db 847 TGGGCTTCTCCTCAGTCATCAAGACTGAAATACAGCAATATTATTCATAGTCATGACAA 906

QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerLeu 40

Db 907 ACTATGAGTGAGACCTTAACAGGCTCCCAAGCAAGAACTATCTTATGAGTCAGTTTA 966

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAspLeuSerAspIleAspGlnAspGlyLysLeu 60

Db 967 CCACAGGCTCAGCTGGCTTCATATGGAATCTTTCTGACATTGATCAAGATGAAACTT 1026

QY 61 ThrAlaGlnGlnPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80

Db 1027 ACAGCAGAGGAATTTATCTCGGCAATGCACTCATTTAGTAGTATGTCGGCAACCA 1086

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArgArg 96

Db 1087 CTGGCACCCTCTCCTGCTCCAGATACATTCACCTCTTTTAGAAGA 1134

RESULT 7
LOCUS BD158570 2131 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD158570
VERSION BD158570.1 GI:27864328
KEYWORDS JP 2002191363-A/13413.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2131)
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 13413 09-JUL-2002;
JOURNAL HELIX RESEARCH INSTITUTE
AUTHORS OS Homo sapiens (human)
COMMENT PN JP 2002191363-A/13413
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10', C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES
source Location/Qualifiers
FT CDS (347)..(2131).
1..2131
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 2.07e-50 Length: 2131
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x BD158570 (1-2131)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20

Db 878 TGGGCTGTTCTCCTCAGTCATCAAGACTGAAATACAGGCAATATTATTCATAGTCATGACAA 937

QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerLeu 40

Db 938 ACTATGAGTGAGACCTTAACAGGCTCCCAAGCAAGAACTATTTCTATGAGTCAGTTTA 997

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAspLeuSerAspIleAspGlnAspGlyLysLeu 60

Db 998 CCACAGGCTCAGCTGGCTTCATATGGAATCTTTCTGACATTGATCAAGATGAAACTT 1057

QY 61 ThrAlaGlnGlnPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80

Db 1058 ACAGCAGAGGAATTTATCTCGGCAATGCACTCATTTAGTAGTATGTCGGCAACCA 1117

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArgArg 96

Db 1118 CTGGCACCCTCTCCTGCTCCAGATACATTCACCTCTTTTAGAAGA 1165

RESULT 8
LOCUS AK027846 2131 bp mRNA linear PRI 01-AUG-2002
DEFINITION Homo sapiens cDNA FLJ14940 fls, clone PLACE1010942, highly similar to Homo sapiens interectin long isoform (ITSN) mRNA.
ACCESSION AK027846
VERSION AK027846.1 GI:14042823
KEYWORDS Oligo capping; fls (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M., Hiraka,S., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuko,Y. and Kanehori,K.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2131)
Isogai,T. and Otsuki,T.
Direct Submission
JOURNAL Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert construction; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
FEATURES
source Location/Qualifiers
1..2131
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_xref="PLACE1010942"
/tissue_type="placenta"
/clone_lib="PLACE1"
/note="Cloning vector: pMT18SFL3"

ORIGIN
Alignment Scores:
Pred. No.: 2.07e-50 Length: 2131
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AK027846 (1-2131)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20

Db 878 TGGGCTGTTCTCCTCAGTCATCAAGACTGAAATACAGGCAATATTATTCATAGTCATGACAA 937

QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerLeu 40

Db 938 ACTATGATGACACTTAAAGCTCCCAAGCAAAATATTTATGACAGCAAGTTA 997
 Qy 41 ProglinalaginLeuA1Ser11eTnpasnuLeuSerAap11eAapG1aAapG1yysleu 60
 Db 998 CCACAGGCTCAGCTGCTTCAATATGAAATCTTTCTGACATTGATCAAGATGAAACTT 1057
 Qy 61 Thr1aagiug1uphe1leuA1aMeth1sleu1leAapVal1a1aMetSerG1yginPro 80
 Db 1058 ACAGAGAGGAATTTATCTGCAATGACACCTGATGATGATGATGCTGGCCAAACA 1117
 Qy 81 LeuProPova1leuProProG1uTyr1leProProSerPheArgArg 96
 Db 1118 CTGCCACCTGCTCTCCCTCCAGAAATACATTCACCTTCTTTAGAGA 1165

RESULT 9
 LOCUS AK074554 2131 bp mRNA linear PRI 03-SEP-2002
 DEFINITION Homo sapiens cDNA FLJ90073 fis, clone HEMBA1004110, highly similar to Homo sapiens interseotin short form mRNA.

ACCESSION AK074554
 VERSION AK074554.1 GI:22760070
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 AUTHORS Isegai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T., Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S., Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Kojima,S., Nagahari,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y., Aotsuka,S., Sasai,N., Hattori,A., Okumura,K., Iwayanagi,T. and Nimomura,K.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2131)
 AUTHORS Isegai,T. and Otsuki,T.
 JOURNAL Direct Submission
 COMMENT Submitted (25-MAR-2002) Takao Isegai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; cDNA 5' & 3' end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).
 FEATURES
 SOURCE Location/Qualifiers
 1..2131
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="HEMBA1004110"
 /tissue_type="whole embryo, mainly head"
 /clone_id="HEMBA1"
 /dev_stage="embryo, 10 weeks"
 /note="Cloning vector: pHE18SFL3"

ORIGIN
 Alignment Scores: 2.07e-50 Length: 2131
 Score: 501.00 Matches: 96
 Percent Similarity: 100.00% Mismatch: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-09-720-934-2_COPY_215_310 (1-96) x AK074554 (1-2131)

Qy 1 TnpalavalProglinalaginSer11eTnpasnuLeuSerAap11eAapG1aAapG1yysleu 20

Db 847 TGGGCTGTCCTCAGTCATCAAGACTGAATATATTCATATGATGACAA 906
 Qy 21 ThmetSerG1yH1sleuThrg1yProglinalaagThr1leuMeG1nSerSerleu 40
 Db 907 ACTATGATGACACTTAAAGCTCCCAAGCAAAATATTTATGACAGCAAGTTA 966
 Qy 41 ProglinalaginLeuA1Ser11eTnpasnuLeuSerAap11eAapG1aAapG1yysleu 60
 Db 967 CCACAGGCTCAGCTGCTTCAATATGAAATCTTTCTGACATTGATCAAGATGAAACTT 1026
 Qy 61 Thr1aagiug1uphe1leuA1aMeth1sleu1leAapVal1a1aMetSerG1yginPro 80
 Db 1027 ACAGAGAGGAATTTATCTGCAATGACACCTGATGATGATGATGCTGGCCAAACA 1086
 Qy 81 LeuProPova1leuProProG1uTyr1leProProSerPheArgArg 96
 Db 1087 CTGCCACCTGCTCTCCCTCCAGAAATACATTCACCTTCTTTAGAGA 1134

RESULT 10
 LOCUS BC058925 2199 bp mRNA linear PRI 05-NOV-2003
 DEFINITION Homo sapiens interseotin 1 (SH3 domain protein), mRNA (cDNA clone IMAGE:4443129), partial cds.

ACCESSION BC058925
 VERSION BC058925.1 GI:37589134
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2199)
 AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,K.H., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stopleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scherz,T.E., Brownstein,M.J., Udell,T.B., Toshyaki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullenb,S.J., Bosak,S.A., McEwan,P.J., McKernan,R.J., Malek,J.A., Gunaratne,P.H., Richards,S., Wooley,K.C., Hale,S., Garcia,A.M., Gay,J.J., Hallyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Heltón,E., Kettleman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Scherchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16893-16903 (2002)
 MEDLINE 22388257
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 2199)
 AUTHORS Strausberg,R.
 JOURNAL Direct Submission
 COMMENT Submitted (01-OCT-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
 COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Genome Sequence Center,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Boeder, Yaron Butlerfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo

Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AF132478 (1-3723)

QY 1 ThrAlaValAlProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
DB 643 TGGGGTGGCTCAGTCTCAAGGCTGAATAACAGCAGTATTCAACAGCACACACAA 702
QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
DB 703 ACTATGATGACACTTAAAGGTCCTCCAGGCAAGACTATTCTCATCAATCAAGTTTA 762
QY 41 ProGlnAlaGlnLeuAlaSerIleThrAsnLeuSerAspIleAspGlnLysLeu 60
DB 763 CCCGAGGCTCAGCTGCTTCAATGAAATCTTCTGCATGATCAAGATGAAACATC 822
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 823 ACTGACAGAAATTTATCTAGTACCACTTAATGATGTTGCCATGTCTGGTCAAGCCA 882
QY 81 LeuProProValLeuProGluTyrIleProProSerPheArgArg 96
DB 883 CTGCCGCGCTCTCTCTCCCAATACATCCCTCTCTCTTCAAGAGA 930

RESULT 12
AF132672 3812 bp mRNA linear ROD 28-JUN-1999
LOCUS Rattus norvegicus EH-domain/SH3-domain containing protein mRNA,
DEFINITION complete cds.
ACCESSION AF132672 GI:4838525
VERSION AF132672.1
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 3812)
Okamoto, M., Schoch, S., and Sudhof, T.C.
EHSN1/intersectin, a protein that contains EH and SH3 domains and
binds to dynamin and SNAP-25. A protein connection between
exocytosis and endocytosis?
J. Biol. Chem. 274 (26), 18446-18454 (1999)

JOURNAL MEDLINE 99303609
PUBMED 10373452
2 (bases 1 to 3812)
Okamoto, M., Schoch, S., and Sudhof, T.C.
Direct Submission
Submitted (02-MAR-1999) Center for Basic Neuroscience and HHMI,
Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX
75235, USA

FEATURES
source 1. 3812 Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
41..3481
/note="EHSN1"
/codon_start=1
/product="EH-domain/SH3-domain containing protein"
/protein_id="ABD31026.1"
/db_xref="GI:4838526"

CDS
1
/translation="MAQPTPTPGSLDIWATVERAKHQFQSLKPISGFITGDOA
RNFPGSLQPVLAQIWLADMNKGMDVEFSIAKLILKIKQGYQLPPLPPVM
KOOPATISAPAFGKGAAMPPLAVAPVPMGSLPVGMSPPLVSVPOAEPPLAN
GAPVIOPLPAFAHRAATLPPSSSRSGPSOLNTXLOKQSPVVASAPAAEWAAP
QSSRIKROLNSHDKTMSGHITGQARTILMOSSLPAQOLASITNLSIDTODGKLT
EEFLIAMLIDIVASGQPLPVLPPEYVPPSPRVRSQSGMSVSSSSADQLPEEPS
SEDEQVKEKKLPVTEDEKRENFERNLELKRRLALDEQKKEERLAQDERAQER

ORIGIN

Alignment Scores:

Pred. No.: 3.9e-50 Length: 3812
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AF132672 (1-3812)

QY 1 ThrAlaValAlProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
DB 683 TGGGGTGGCTCAGTCTCAAGGCTGAATAACAGCAGTATTCAACAGTCAAGCAAG 742
QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
DB 743 ACCATGATGACACTTAAAGGTCCTCCAGGCAAGACTATTCTCATCAATCAAGTTTA 802
QY 41 ProGlnAlaGlnLeuAlaSerIleThrAsnLeuSerAspIleAspGlnLysLeu 60
DB 803 CCCGAGGCTCAGCTGCTTCAATGAAATCTTCTGCATGATCAAGATGAAACATC 862
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 863 ACCGACAGAAATTTATCTAGTACCACTTAATGATGTTGCCATGTCTGGTCAAGCCA 922
QY 81 LeuProProValLeuProGluTyrIleProProSerPheArgArg 96
DB 923 CTGCCGCGCTCTCTCTCCCAATACATCCCTCTCTCTTCAAGAGA 970

RESULT 13

AF127798 4025 bp mRNA linear ROD 28-JUN-1999
LOCUS Rattus norvegicus EH- and SH3-domain containing protein EHSN1 mRNA,
DEFINITION complete cds.
ACCESSION AF127798
VERSION AF127798.1 GI:4835852
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 4025)
Okamoto, M., Schoch, S., and Sudhof, T.C.
EHSN1/intersectin, a protein that contains EH and SH3 domains and
binds to dynamin and SNAP-25. A protein connection between
exocytosis and endocytosis?
J. Biol. Chem. 274 (26), 18446-18454 (1999)

JOURNAL MEDLINE 99303609
PUBMED 10373452
2 (bases 1 to 4025)
Okamoto, M., Schoch, S., and Sudhof, T.C.
Direct Submission
Submitted (11-FEB-1999) Center for Basic Neuroscience and HHMI,
Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX
75235, USA

FEATURES

Source

CDS

[illegible]

ORIGIN

Alignment Scores:

Pred. No.:	4,14e-50	4025
Score:	501.00	96
Percent Similarity:	100.00%	0
Best Local Similarity:	100.00%	0
Query Match:	100.00%	0
DB:	10	0
	Gaps:	0

QY	TYPALVAALProGlnSerArgLeuLysTyrAGGGLneupePasnSerHisAspIys	20
Db	683 TGGGTCTGTGCCTACGTCGTCMAAGACTGAATCAGGGCATTAATTCACAGCAACAAG	742
QY	21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgHrileLeuMetGlnSerSerleu	40
Db	743 ACCATAGTGACACCTTAACAGGTCCCGGCAAAATATTTCATCATCAAGTTTA	802
QY	41 ProGlnAlaGlnLeuAlaSerIleTetrapeuSenSarpIIeaSpGInaSpGlyLytleu	60
Db	803 CCCAAGCTAGCTGGCTTCATACTGAAATCTTTCGACATTGATCAAGAAGAAAAGCTC	862
QY	61 ThraIagLuGIpheIleuAlaMetHisIleuIIeaSPValAlaMeSerGlyGlnPro	80
Db	863 ACCGCAAGAAATTCATCTTAGGATGACACCGATTGATGTGGATATGTCTGGTCAACCA	922
QY	81 LeuProProValleuProProGluIyrlleProProSerPheIaGhg	96
Db	923 CTTCGGCTGTCTGTGCCTCAAAATAACATCCCTCTTCTTTAANAAG	970

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 4321)	Labber, J., Bahr, A., Mewes, H. W., Well, B., Amid, C., Oeanger, A., Fobo, G., Han, M. and Wiemann, S.	Direct Submission	Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuheideberg, GERMANY	
			Clone from S. Wiemann, Molecular Genome Analysis, German Cancer	

FEATURES

Source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="21q22.11"
/clone="DKFZp686J17173"
/tissue_type="human uterus endometrial primary cell culture"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_SfiI, host
DH10B; sites SfiI + SfiIB"
/dev_stage="adult"
1..4321
/gene="DKFZp686J17173"
<1..1886
gene
DSD

```

ORIGIN

```
polyA_signal      4267.  .4272  
                  /gene="DKFZp686J17173"  
polyA_site        4289  
                  /gene="DKFZp686J17173"
```

Alignment Scores:

Pred. No.:	4,47e-50	length:	4322
Score:	501.00	Matches:	96
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

[illegible]

```
Qy 41 ProGln1aGln1eua1aSer1eTTPaSn1euserAasp1leaSpG1nAspG1yLys1eu 60
Db 330 CCACAGGCTCAGCTGGCTTCATATGGAATCTTTCTGCACTTGAATGAGATGAGAAACTT 389
Qy 61 Thr1aGln1eua1aSer1eTTPaSn1euserAasp1leaSpG1nAspG1yLys1eu 80
Db 390 ACACAGAGGAAATTTATCTGCTGCACTTGAATGAGATGAGATGAGAAACTT 449
Qy 81 LeuProPVal1euaProProG1uTyr1leProProSerPhe1aG1y 96
Db 450 CTGCACCTGCTCCTCCCTCCAGAAATACATCCACCTCTTTAAGAAGA 497

RESULT 15
AF132481 5145 bp mRNA linear ROD 09-MAR-1999
LOCUS AF132481
DEFINITION Mus musculus Esei1 protein mRNA, complete cds.
ACCESSION AF132481
VERSION AF132481.1 GI:4378890
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 5145)
Sengar, A.S., Wang, W., Bishay, J., Cohen, S. and Egan, S.E.
The EH and SH3 domain Ese proteins regulate endocytosis by linking
to dynamin and Eps15
EMBO J. 18 (5), 1159-1172 (1999)
JOURNAL MEDLINE
99164083
PUBMED 10064583
REFERENCE 2 (bases 1 to 5145)
Sengar, A.S., Wang, W., Cohen, S., Bishay, J. and Egan, S.E.
Direct Submission Programs in Cancer & Blood
Submitted (02-MAR-1999) Research/Developmental Biology, The Hospital for Sick Children, 555
University Avenue, Toronto, ON M5G-1X8, Canada
JOURNAL Location/Qualifiers
1..5145
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..5145
/codon_start=1
/product="Esei1 protein"
/protein_id="AA019749.1"
/db_xref="GI:4378891"
/translation="MAQPTPTFGSLDVAITVERAKHQDQFLSKPIAGFTTGDA
RNFFFGSLPQVLAQIWAIDMNNNDMDQVESIAMKLIKLIQGYLPSTLPVW
KQOVALISSAPAFGIGIASMPPLTAAPVPMGSIPIVGMSPPLVSSVPAAPVLA
GAPVIOPLPAFAHPAATWPKSSFSRSGSOLNTKLOKASPDVVASPPAAEMAV
OSSLKYKQIPNSHDKTMSGHITGPOARTTMMOSISIPOLASITMNLSDIDQKRTA
EFTLAHLIDVHNSGQPLPVLPPEYITPPSFRVRSQMSVISSSVDORLPBES
SEDEQOPEKKLPVTFEDKKRNPFGSVLEKROALDQQRKEORLEAQLERAEQR
KERROBOEAKROLEKOLEKORELEBERREKLETEREAARELEORLEORLE
NRROELNQRNKEQBJEVLKARKKTLFELEALNDKKHOLEGKLODIRCLAROLE
IESTKSRLEILAEITTHLOOLOSOQMLGRIPEKOLISPOLKVOONSHPRDSILT
LKRLAEKELAROLRBDLEVERETRSKGLRIDVENNOLKELREIHSKQOLOKORSL
EALRLKQEQERKSELEKQEKDARVQOERDKOMLEHYQOEBQPRPKREDEBRK
EDSVRKKEAREARWODKOSRLPHQBPAKLTOAPWSTTEKGLTISAQBSKV
VYVALYPPESRSHDEITOPDIYHVEDSQTGEVGMGLKGTGMFPANVAKIP
ENEVPTPAKPTDLSAPAKLARETAPLPVTSSESTPNWADFSSTWPSNSN
KEPTNDMDTMAAOPSLTVSAGOLRORSAFPATATGSSPSVLGQGEVGLQOAL
YPMRAKKNHINFNKSDVITVLEBODMMWPGVOGKGMFKSVYKLSGPRKSTSI
DTGPTESPASLKRVASPAKPAIPEEETIAMTTTSSGCGITLTOGQDVITYTKKDG
WMTGTVDKSGVPPSNVRLKDSGSGTAGTSGKPEIAQVIASTAAATGPEQLT
APGQILIRKKNPGWMEGLQARGKROIGMFPANVYKLSPGSKITPELPTAV
OPAVCOVITGAYDAONDDDLAFSGQIINVLNKEDPMWKGESVGGGLPSPNYKL
TTDMDPGQWCSDLTLDMLTPTERKRGYTHELITVENYVNDLOLTVBEI FORKLE
SELTLEKVMATFYWKELIMCNTKILKALVRKMGSEKMPVMIGDIISAOLPHMO
PYIRFCSCOLGAALIOOKTDBADPFKEVFRLANDPRCKMPLSSFLIKPMORYRY
PLITKNILENTPENRPHSHLHALEKAEELCSQVNEGVRKENSRLIEMIQAHVQCE
GLSEOLVENSVTNCLSPRKFLHSGKLYAKSKNKEIYGLFNDFFLLTOITKPLGSSGT
```

```
ORIGIN
Alignment Scores:
Pred. No.: 5,41e-50 Length: 5145
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AF132481 (1-5145)
Qy 1 TTPAlaVa1ProGln1euserAasp1leaSpG1nAspG1yLys1eu 20
Db 643 TGGGCTGCTCAGTATCAAGGCTGMAATACAGGCGAGTATTCACACGACAGCAA 702
Qy 21 ThrMetSerG1yH1sleuThrG1yProGln1aArGThr1leuMerGlnSerSer1eu 40
Db 703 ACTATGAGTGACACTTAACAGGTCCCGACAGAACTATTCATCATCAAGTTTA 762
Qy 41 ProGln1aGln1eua1aSer1eTTPaSn1euserAasp1leaSpG1nAspG1yLys1eu 60
Db 763 CCCACAGGCTCAGCTGGCTTCATATGGAATCTTTCTGCACTTGAATGAGATGAGAAACTT 822
Qy 61 Thr1aGln1eua1aSer1eTTPaSn1euserAasp1leaSpG1nAspG1yLys1eu 80
Db 823 ACTCAGAGAAATTTATCTGCTGCACTTGAATGAGATGAGATGAGAAACTT 882
Qy 81 LeuProPVal1euaProProG1uTyr1leProProSerPhe1aG1y 96
Db 883 CTGCACCTGCTCCTCCCTCCAGAAATACATCCCTCTTTAAGAAGA 930
```

Search completed: July 1, 2004, 16:14:11
Job time : 2196.79 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 11:41:07 ; Search time 296.319 Seconds
(without alignments)
1376.313 Million cell updates/sec

Title: US-09-720-934-2_COPY_215_310
Perfect score: 501
Sequence: 1 WAVPOSSRLKXROLFNSHDK.....SQQLPVPVLPPEYIPPSFRR 96

Scoring table:
BL0SUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-O/cp2n2.1/USPTO.spool_p/US09720934/runat.30062004.064539.13422/app.query.fasta_1.1386
-DB=N Geneseg 297and4 -QPM=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US09720934 @CGM 1.1 1596 @runat.30062004.064539.13422 -NCPU=6 -ICPU=3
-NO MMAP -LARGEJITTERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseg 297and4.*

1: geneseg1980s.*
2: geneseg1990s.*
3: geneseg2000s.*
4: geneseg2001as.*
5: geneseg2001bs.*
6: geneseg2002s.*
7: geneseg2003as.*
8: geneseg2003bs.*
9: geneseg2003cs.*
10: geneseg2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	501	100.0	1676	4	AAK94611	AAK94611 Human CDN
2	501	100.0	2079	2	AAZ34573	Aaz34573 Human SH3
3	501	100.0	2131	4	AAK94139	Aak94139 Human ful
4	501	100.0	2131	4	AAH16578	AaH16578 Human CDN
5	501	100.0	3723	3	AAZ39009	Aaz39009 Mouse Ese
6	501	100.0	5082	3	AAZ39008	Aaz39008 Mouse Ese
7	501	100.0	5144	3	AAZ39025	Aaz39025 Mouse Ese
8	501	100.0	5195	2	AAZ34572	Aaz34572 Human SH3

9	501	100.0	5199	2	AAZ34570	Aaz34570 Human SH3
10	501	100.0	5458	2	AAZ34571	Aaz34571 Human SH3
11	501	100.0	5738	3	AAZ39024	Aaz39024 Mouse Ese
12	496	99.0	548	4	AA180000	Ab180000 Human pol
13	489	97.6	3319	4	ABK43498	AbK43498 DNA encod
14	489	97.6	3466	4	AA163825	Aa163825 Human pol
15	416	83.0	7435	5	AA584763	AA584763 DNA encod
16	349	69.7	2017	3	AAA69762	AAA69762 Human ova
17	349	69.7	2017	3	ABN72656	ABn72656 Ovarian c
18	349	69.7	2017	6	ADA08821	Ada08821 Human ova
19	349	69.7	3593	3	AAZ39011	Aaz39011 Mouse Ese
20	349	69.7	4625	3	AAZ39010	Aaz39010 Mouse Ese
21	349	69.7	4975	3	AAZ39027	Aaz39027 Mouse Ese
22	349	69.7	6014	3	AAZ39026	Aaz39026 Mouse Ese
23	345	68.9	4447	4	AA502055	AA502055 DNA encod
24	345	68.9	5828	6	AA147247	AA147247 Allergic
25	345	68.9	6103	4	AAK52332	Aak52332 Human pol
26	337	67.3	346	3	AA444038	Aa444038 Human sec
27	334	66.7	955	4	AA163826	Aa163826 Human pol
28	334	66.7	955	4	AA531620	AA531620 DNA encod
29	334	66.7	955	4	ABK43501	AbK43501 DNA encod
30	292	58.3	3981	4	AB101995	Ab101995 Drosophila
31	292	58.3	7225	4	AB101994	Ab101994 Drosophila
32	280	55.9	831	4	AAH03435	AaH03435 Human CDN
33	280	55.9	1329	4	AAH15260	AaH15260 Human CDN
34	208	41.5	480	3	AA69750	AA69750 Human ova
35	208	41.5	480	6	ABN72644	ABn72644 Ovarian c
36	208	41.5	480	8	ADA08809	Ada08809 Human ova
37	208	41.5	531	3	AAA69695	AAa69695 Human ova
38	208	41.5	531	6	ABN72589	ABn72589 Ovarian c
39	208	41.5	531	8	ADA08754	Ada08754 Human ova
40	204.5	40.8	3033	2	AAV16484	AaV16484 CDNA enco
41	204.5	40.8	3033	2	AAV13999	AaV13999 Murine ep
42	204.5	40.8	3033	2	AAH04192	AaH04192 Murine ep
43	203.5	40.6	2666	4	AAH18453	AaH18453 Human CDN
44	203.5	40.6	2774	3	AAZ96752	Aaz96752 Nuclear t
45	203.5	40.6	2782	4	AAH19248	AaH19248 Human sec

ALIGNMENTS

RESULT 1
AAK94611 standard; CDNA; 1676 BP.
ID AAK94611;
AC AAK94611;
XX
XX
DT 07-NOV-2001 (first entry)
XX
XX
DE Human CDNA, SEQ ID NO: 3565.
XX
XX
KW Human, full length cDNA; CDNA synthesis; oligo-capping; ss.
XX
XX
OS Homo sapiens.
XX
XX
PN EP1130094-A2.
XX
XX
PD 05-SEP-2001.
XX
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
XX
PA (HELI-) HELIX RES INST.
XX
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX
WP1; 2001-524255/58.
XX
XX
P-PSDB; AAM93676.
XX
XX

PT 830 Primers useful for synthesizing full length cDNA clones and their use
in genetic manipulation.
XX
XX
PS Disclosure; SEQ ID NO 3565; 1380bp + Sequence Listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
clones. 830 cDNA molecules encoding a human protein have been isolated
and nucleotide sequences of 5' and 3' ends of the cDNA molecules have
been determined. Primers for synthesizing the full length cDNA are useful
for clarifying the function of the protein encoded by the cDNA. The full
length clones were obtained by construction of full length enriched cDNA
libraries that were synthesized by the oligo-capping method. The primers
enable the production of the full length cDNA easily without any special
methods. The present sequence is a human cDNA provided in the
specification. Note: The sequence data for this patent did not form part
of the printed specification, but was obtained in CD-ROM format directly
from EPO
XX
SQ Sequence 1676 BP; 485 A; 381 C; 458 G; 352 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,02e-52 Length: 1676
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-720-934-2_COPY_215_310 (1-96) x AA394611 (1-1676)
QY 1 TTPALAVLProGInSeSerArgLeuLysTYrArgGInLeuPheAsnSerHisAspLys 20
Db 906 TGGGCTGTTCCTCACTCATCAAGACTGAATAACAGGCAATTATTCATATGATCATACAA 965
QY 21 ThrMetSerGlyHisLeuThyGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 966 ACTATGATGAGACACTTAAACAGCTCCCAAGAGAAACTATCTTATCAGTCAAGTTTA 1025
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
Db 1026 CCACAGGCTTCAGCTGCTTCATATGAAATCTTCTTGACATTGATCAAGATGAAACTT 1085
QY 61 ThrAlaGlnIupheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 1086 ACAGAGAGGAAATTATTCCTGCAATGACCTCATTTGATGATGATCTGCGCAACCA 1145
QY 81 LeuProProValLeuProProGlnIuTyrIleProProSerPheArgArg 96
Db 1146 CTGGCACCCTGCTGCTCCAGAAATACATTCACCTCTTTAGAGAGA 1193
RESULT 2
AA234573
ID AA234573 standard; cDNA; 2079 BP.
XX
AC AA234573;
XX
DT 01-FEB-2000 (first entry)
XX
XX Human SH3D1A cDNA clone 5.
XX
KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
KW megakaryocytic abnormality; myeloproliferative disorder;
KW platelet disorder; neural disorder; thrombocytopenia;
KW haematopoietic disorder; cognitive dysfunction; microcephaly;
KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 136..2079
FT CDS
FT /tag= a
XX
XX PN W09953062-A2.

XX
PD 21-OCT-1999.
XX
XX
PF 16-APR-1999; 99WO-US008371.
XX
PR 16-APR-1999; 98US-0082007P.
XX
PA (CEDA-) CEPARS SINAI HEALTH SYSTEM.
XX
PI Korenberg JR, Chen X;
XX
XX WPI; 1999-633829/54.
DR P-PsDB; AA32157.
XX
PT Nucleic acid from the human SH3D1A gene and its products, useful for the
diagnosis and treatment of myeloproliferative disorders and leukemia.
XX
PS Claim 2; Fig 12; 99pp; English.

CC This is the nucleotide sequence of a cDNA clone, termed clone 5,
CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
CC contributes to the development of platelets and the pathogenesis of
CC leukaemias, both in general and in particular those involving the
CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
CC region for low platelets on chromosome 21. Sequencing of 5 different
CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
CC least 3 isoforms exist. The invention provides methods for the diagnosis
CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
CC association of gains in chromosome 21 with leukaemias, neural
CC abnormalities, dysfunctions and disorders including brain malformations
CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
CC and colpocephaly. Methods are also provided for: suppressing cells unable
CC to regulate themselves; screening for a somatic alteration in the SH3D1A
CC gene; monitoring the progress and adequacy of a treatment; monitoring
CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and
CC treatment of a subject (including a prenatal subject) having
CC megakaryocytic abnormality, myeloproliferative disorder, platelet
CC disorder, leukaemia or neural disorder using a nucleic acid that
CC expresses SH3D1A or its antisense nucleic acid
XX

SO Sequence 2079 BP; 700 A; 451 C; 495 G; 433 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.37e-52 Length: 2079
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AA234573 (1-2079)

QY 1 TTPALAVLProGInSeSerArgLeuLysTYrArgGInLeuPheAsnSerHisAspLys 20
Db 778 TGGGCTGTTCCTCACTCATCAAGACTGAATAACAGGCAATTATTCATATGATCATACAA 837
QY 21 ThrMetSerGlyHisLeuThyGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 838 ACTATGATGAGACACTTAAACAGCTCCCAAGAGAAACTATCTTATGCAATCAAGTTTA 897
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
Db 898 CCACAGGCTCAGCTGCTTCATATGAAATCTTCTGACATTGATGATCAAGATGAAACTT 957
QY 61 ThrAlaGlnIupheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 958 ACAGAGAGGAAATTATTCCTGCAATGACCTCATTTATGATGATGATCTGCGCAACCA 1017
QY 81 LeuProProValLeuProProGlnIuTyrIleProProSerPheArgArg 96

Db 1018 CTGGCACCCTGCTCCTCCAGAAATACATTCACCTTCTTTAGAGA 1065

RESULT 3

AAK94139

ID AAK94139 standard; cDNA; 2131 BP.

XX

XX AAK94139;

AC

XX 06-NOV-2001 (first entry)

DT

XX

DE Human full-length cDNA, SEQ ID NO: 2646.

XX

XX Human, full length cDNA; cDNA synthesis; oligo-capping; ss.

KM

XX Homo sapiens.

OS

XX EP1130094-A2.

PN

XX 05-SEP-2001.

PD

XX 07-JUL-2000; 2000EP-00114089.

PF

XX 08-JUL-1999; 99JP-00194486.

PR 11-JAN-2000; 2000JP-00118774.

PR 02-MAY-2000; 2000JP-00183765.

XX

XX (HELI-) HELIX RES INST.

PA

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX

DR WPI; 2001-524255/58.

DR P-PSDB; AAM93229.

XX

PT 830 Primers useful for synthesizing full length cDNA clones and their use

PT in genetic manipulation.

XX

XX

PS Claim 8; SEQ ID NO 2646; 1380bp + Sequence Listing; English.

XX

CC The invention relates to primers for synthesizing full length cDNA

CC clones. 830 cDNA molecules encoding a human protein have been isolated

CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have

CC been determined. Primers for synthesizing the full length cDNA are useful

CC for clarifying the function of the protein encoded by the cDNA. The full

CC length clones were obtained by construction of full length enriched cDNA

CC libraries that were synthesised by the oligo-capping method. The primers

CC enable the production of the full length cDNA easily without any special

CC methods. The present sequence is a full length human cDNA of the

CC invention. Note: The sequence data for this patent did not form part of

CC the printed specification, but was obtained in CD-ROM format directly

CC from EPO

XX

SQ Sequence 2131 BP; 694 A; 469 C; 526 G; 442 T; 0 U; 0 Other;

XX

Alignment Scores:

Pred. No.: 1.41e-52 Length: 2131

Score: 501.00 Matches: 96

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 4 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AAK94139 (1-2131)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20

Db 847 TGGGCTGTCCTCCTCATCATCAAGACTGAATACAGGCAATTATTCAATAGTCATGACAA 906

QY 21 ThwTserGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIleu 40

Db 907 ACTATGAGTGGACACTTAACAGGTCGCCAAGCAAGAACTATTCTTATGCGATCAAGTTTA 966

QY 41 ProGlnIaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60

Db 967 CCAAGGCTCAGCTGGCTTAAATATGAACTTTCTGACTGATCAAGATGAAAACCTT 1026

QY 61 ThrAlaGlnIubPheIleuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80

Db 1027 ACAGCAGAGAAATTTATCTCGGCAATGCACCTCATTTAGTATGTCGTGGCAACCA 1086

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArgArg 96

Db 1087 CTGGCACCCTGCTCCTCCAGAAATACATTCACCTTCTTTAGAGA 1134

RESULT 4

AAH16578

ID AAH16578 standard; cDNA; 2131 BP.

XX

XX AAH16578;

AC

XX 26-JUN-2001 (first entry)

DT

XX

DE Human cDNA sequence SEQ ID NO:16658.

XX

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

KM

XX Homo sapiens.

OS

XX EP1074617-A2.

PN

XX 07-FEB-2001.

PD

XX 28-JUL-2000; 2000EP-00116126.

PF

XX 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX

XX (HELI-) HELIX RES INST.

PA

PI Ota T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Makamatsu A, Nagai K, Otsuki T;

XX

DR WPI; 2001-318749/34.

XX

XX

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-

PT length cDNAs defined in the specification, and for the detection and/or

PT diagnosis of the abnormality of the proteins encoded by the full-length

PT cDNAs.

XX

PS Claim 8; SEQ ID NO 16658; 2537pp + Sequence Listing; English.

XX

XX The present invention describes primer sets for synthesizing 5602 full-

CC length cDNAs defined in the specification. Where a primer set comprises:

CC (a) an oligo-dr primer and an oligonucleotide complementary to the

CC complementary strand of a polynucleotide which comprises one of the 5602

CC nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in the

CC specification. The primer sets can be used in antisense therapy and in

CC gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH16628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893

CC represent human amino acid sequences; and AAH13629 to AAH13632 represent

CC oligonucleotides, all of which are used in the exemplification of the

CC present invention

XX SQ Sequence 2131 BP; 691 A; 448 C; 572 G; 420 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	2131
Score:	501.00	Matches: 96
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	4	Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AAH16578 (1-2131)

QY 1 TrpAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20

DB 878 TGGGCTGTTCCTCACTGATCAAGACTGAATACAGGCAATTTATTAATGATGATACAAA 937

QY 21 ThrmSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40

DB 938 ACTATGATGAGGACCTTAACAGGTCCCAAGCAAGAACATTTCTATGAGTCAAGTTTA 997

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPasnLeuSerAspIleAspGlnAspGlyLysIeu 60

DB 998 CCACAGGCTCAGCTGGCTTCATATGAAATCTTCTGACATTGATCAAGATGGAACCTT 1057

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80

DB 1058 ACAGAGAGAGAAATTTATCTCTGCAATGCACTCATTTGATGATGCTGAGCCACCA 1117

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArgArg 96

DB 1118 CTGGCCACTGTCCTCTCCCAATATACATTCACCTTCTTTAGAAGA 1165

RESULT 5

AAZ39009

ID AAZ39009 standard; cDNA, 3723 BP.

XX AC AAZ39009;

XX DT 28-FEB-2000 (first entry)

XX DE Mouse Esei coding sequence.

XX KW Mouse; murine; Esei; Ees2; endocytosis; vesicular trafficking; regulation; actin cytoskeleton; detection; cancer; infection; EH-domain and SH3-domain regulator of endocytosis; anticancer; antiproliferative; antiviral; ss.

XX OS Mus sp.

XX PN WO9955728-A2.

XX PD 04-NOV-1999.

XX PF 27-APR-1999; 99WO-CA000375.

XX PR 27-APR-1998; 98CA-02230201.

XX PR 05-FEB-1999; 99US-0118739P.

XX PA (HSCR-) HSC RBS & DEV LP.

XX PI Egan SE, Wang W, Sengar A;

XX DR WPI: 2000-052802/04.

XX DR P-PSDB; AAY57444.

XX PT New nucleic acid encoding Esei and 2 proteins, involved in regulation of endocytosis, used e.g. for treating cancer or preventing viral infection.

XX PS Claim 6; Page 40-42; 99pp; English.

XX CC The present sequence encodes mouse Esei. The present invention specifically describes mammalian Esei and 2 proteins (I) and their splice

CC variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (I) CC are involved in regulation of clathrin-mediated endocytosis (as a complex CC with Ees1s protein), vesicular trafficking and actin cytoskeleton. CC Generally (I) (or its (ant)agonists, mimetics, fragments and inactive CC mutants); (I)-specific antibodies (Ab); sequences antisense to the (I) CC polynucleotide; agents that downregulate expression of Ese genes or CC antagonists of an Ese binding partner are used to treat diseases CC associated with undesirable endocytosis and resulting changes in cellular function. Particularly overexpression of Esei is used to block clathrin-mediated endocytosis in vivo or in cell cultures, while administration of CC (I) is used to promote endocytosis of selected cells. (Ant)agonists of CC (I) or Ab are used to suppress abnormal proliferation of cells that can be stimulated to proliferate by a growth factor receptor; and similar CC compounds (also inactive Ese mutants) can be used to prevent viral infection. Endocytosis may also be regulated, in vivo or in cell cultures, by forming an Ese-Ees1s complex, then binding dynam to the CC complex. Generally conditions that can be treated include cancer; CC abnormal cell division or migration; viral infection; or abnormal CC receptor signaling, tissue development or synaptic transmission

XX SQ Sequence 3723 BP; 1065 A; 941 C; 1025 G; 692 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	3723
Score:	2.97e-52	Matches: 96
Percent Similarity:	501.00	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	3	Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AAZ39009 (1-3723)

QY 1 TrpAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20

DB 643 TGGGCTGTTCCTCACTGATCAAGGCTGAATACAGGCAATTTATTAACAGCCACACAAA 702

QY 21 ThrmSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40

DB 703 ACTATGATGAGGACCTTAACAGGTCCCAAGCAAGAACATTTCTATGAGTCAAGTTTA 762

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPasnLeuSerAspIleAspGlnAspGlyLysIeu 60

DB 763 CCCACAGGCTCAGCTGGCTTCATATGAAATCTTCTGACATTGATCAAGATGGAACCTT 822

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80

DB 823 ACTGAGAGAGAAATTTATCTCTGCAATGCACTCATTTGATGATGCTGAGCCACCA 882

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArgArg 96

DB 883 CTGGCCGCCGTCCTCCCAATATACATTCCTCTCTTCAGAGA 930

RESULT 6

AAZ39008

ID AAZ39008 standard; cDNA, 5082 BP.

XX AC AAZ39008;

XX DT 28-FEB-2000 (first entry)

XX DE Mouse Esei full length cDNA sequence.

XX KW Mouse; murine; Esei; Ees2; endocytosis; vesicular trafficking; regulation; actin cytoskeleton; detection; cancer; infection; EH-domain and SH3-domain regulator of endocytosis; anticancer; antiproliferative; antiviral; ss.

XX OS Mus sp.

XX PN WO9955728-A2.

XX PD 04-NOV-1999.

```

PF 27-APR-1999; 99WO-CA000375.
XX
PR 27-APR-1998; 98CA-02230201.
PR 05-FEB-1999; 99US-0118739P.
XX
PA (HSCR-) HSC RES & DEV LP.
XX
PI Egan SE, Wang W, Sengar A;
XX
DR WPI; 2000-052802/04.
DR P-PSDB; AAY57444.
XX
PT New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.
XX
PS Claim 6; Page 38-40; 99pp; English.
XX
XX The present sequence encodes mouse Ese1. The present invention
CC specifically describes mammalian Ese1 and 2 proteins (I) and their splice
CC variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (I)
CC are involved in regulation of clathrin-mediated endocytosis (as a complex
CC with Esp15 protein), vesicular trafficking and actin cytoskeleton.
CC Generally (I) (or its (ant)agonists, mimetics, fragments and inactive
CC mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
CC polynucleotide; agents that downregulate expression of Ese genes or
CC antagonists of an Ese binding partner are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of Ese1 is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (I) is used to promote endocytosis of selected cells. (Ant)agonists of
CC (I) or Ab are used to suppress abnormal proliferation of cells that can
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-Esp15 complex, then binding dynamin to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission
XX
SQ Sequence 5082 BP; 1363 A; 1253 C; 1391 G; 1075 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 4.5e-52 Length: 5082
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-720-934-2_COPY_215_310 (1-96) x AAZ39008 (1-5082)
QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAenSerHisAspLys 20
Db 901 TGGGCTGTGCTCAGTCATCAAGCTCAATATACAGGAGATTATTCAACACGACGACAA 960
QY 21 ThMeSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerLeu 40
Db 961 ACTATGATGAGACACTTAACAGGCTCCAGGACGAAAGAACTATTCACAGCAATCAAGTTTA 1020
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
Db 1021 CCCAGGCTCAGCTGGTTCAATATGGAATCTTTTCACATTGATCAAGATGAAAACTC 1080
QY 61 ThrAlaGlnGlnPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 1081 ACTGCAAGAAAGATTATCTCAGCTATGCACTTAATGATGTGCGCATGCTGAGCGCA 1140
QY 81 LeuProProValLeuProProGlnTyrIleProProSerPheArgArg 96
Db 1141 CTGGCGCCGCTGCTGCTCCAGAAATACATCCCTCTCTTCAGAAAGA 1188
RESULT 7
AAZ39025

```

```

ID AAZ39025 standard; cDNA; 5144 BP.
XX
AC AAZ39025;
XX
DT 28-FEB-2000 (first entry)
XX
DE Mouse Ese1l coding sequence.
XX
KW Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;
KW regulation; actin cytoskeleton; detection; cancer; infection;
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
KW antiproliferative; antiviral; ss.
XX
OS Mus sp.
XX
PN WO9955728-A2.
XX
PD 04-NOV-1999.
XX
PF 27-APR-1999; 99WO-CA000375.
XX
PR 27-APR-1998; 98CA-02230201.
PR 05-FEB-1999; 99US-0118739P.
XX
PA (HSCR-) HSC RES & DEV LP.
XX
PI Egan SE, Wang W, Sengar A;
XX
DR WPI; 2000-052802/04.
DR P-PSDB; AAY57449.
XX
PT New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.
XX
PS Claim 6; Page 59-62; 99pp; English.
XX
XX The present invention specifically describes mammalian Ese1 and 2
CC proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
CC regulator of endocytosis). (I) are involved in regulation of clathrin-
CC mediated endocytosis (as a complex with Esp15 protein), vesicular
CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
CC mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
CC sequences antisense to the (I) polynucleotide; agents that downregulate
CC expression of Ese genes or antagonists of an Ese binding partner are used
CC to treat diseases associated with undesirable endocytosis and resulting
CC changes in cellular function. Particularly overexpression of Ese1 is used
CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
CC administration of (I) is used to promote endocytosis of selected cells.
CC (Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
CC cells that can be stimulated to proliferate by a growth factor receptor;
CC and similar compounds (also inactive Ese mutants) can be used to prevent
CC viral infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-Esp15 complex, then binding dynamin to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission. The
CC present sequence represents mouse Ese1l coding sequence
XX
SQ Sequence 5144 BP; 1447 A; 1298 C; 1415 G; 984 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 4.58e-52 Length: 5144
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-720-934-2_COPY_215_310 (1-96) x AAZ39025 (1-5144)
QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAenSerHisAspLys 20
Db 643 TGGGCTGTGCTCAGTCATCAAGCTCAATATACAGGAGATTATTCAACACGACGACAA 702

```

QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
 DB 703 ACTATGAGTGGACACTTAAAGGTCCCGAGCAAACTATTCTCATCAATCAAGTTTA 762
 QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPasnLeuSerAspIleAspGlnAspGlyLysIeu 60
 DB 763 CCCGAGGCTGAGCTGGCTTCATATGAGAACTCTTTCTGACATTGATCAAGATGGAATACTC 822
 QY 61 ThrAlaGluGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
 DB 823 ACTGAGAGAAATTTATCTAGCTATGACCACTAATTGATGTGCATGTCTGGTCAGCA 882
 QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArg 96
 DB 883 CTGCGCGCCGTCTGCTCCAGAAATACATCCCTCTCTTTCAGAGA 930
 RESULT 8
 AAZ34572 ID AAZ34572 standard; cDNA; 5195 BP.
 AC AAZ34572;
 XX
 XX 01-FEB-2000 (first entry)
 DT
 XX Human SH3D1A cDNA clone 11.
 DE
 XX SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 239..3886
 FT /*tag= a
 XX
 XX MO9953062-A2.
 PN
 XX 21-OCT-1999.
 PD
 XX 16-APR-1999; 99WO-US008371.
 PF
 XX 16-APR-1998; 98US-0082007P.
 PR
 XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 PA
 XX Korenberg JR, Chen X;
 PI
 DR WPI, 1999-633829/54.
 DR P-PSDB; AAY32156.
 DR
 XX
 XX Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 PS Claim 2; Fig 10; 99pp; English.
 XX
 XX This is the nucleotide sequence of full-length cDNA (clone 11)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukaemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AAZ34570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC platelet disorder on chromosome 21, low platelets in deletion for 21,
 CC association of gains in chromosome 21 with leukaemias, neural
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,

CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukaemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 XX
 SQ Sequence 5195 BP; 1592 A; 1148 C; 1237 G; 1218 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4,64e-52 Length: 5195
 Score: 501.00 Matches: 96
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 2
 US-09-720-934-2_COPY_215_310 (1-96) x AAZ34572 (1-5195)
 QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
 DB 881 TGGGCTGTCTCTCAGTCATCAAGACTGAATACAGGCAATATTCAATAGTCATGACAAA 940
 QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
 DB 941 ACTATGAGTGGACACTTAAAGGTCCCGAGCAAACTATTCTTATCAAGTCAAGTTTA 1000
 QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPasnLeuSerAspIleAspGlnAspGlyLysIeu 60
 DB 1001 CCACAGGCTCAGCTGGCTTCATATGAGAACTCTTCTGACATTGATCAAGATGGAATACTT 1060
 QY 61 ThrAlaGluGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
 DB 1061 ACAGAGAGAAATTTATCTGCGCAATGACCTCATTTGATGTAGCTATGTGGCCACACA 1120
 QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArg 96
 DB 1121 CTGCCACTGTCTGCTCCAGAAATACATTCACCTCTTTCAGAGA 1168
 RESULT 9
 AAZ34570 ID AAZ34570 standard; cDNA; 5199 BP.
 AC AAZ34570;
 XX
 XX 01-FEB-2000 (first entry)
 DT
 XX Human SH3D1A cDNA.
 DE
 XX
 XX SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 208..3642
 FT /*tag= a
 XX
 XX MO9953062-A2.
 PN
 XX 21-OCT-1999.
 PD
 XX 16-APR-1999; 99WO-US008371.
 PF
 XX 16-APR-1998; 98US-0082007P.
 PR
 XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 PA


```
Db      909 TGGGCTGTCTCTAGTCAATCAAGTGAATACAGGCAATTATTCATATGACACACAAA 968
Qy      21 ThwterSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerIeu 40
Db      969 ACTATGAGTGACACTTACAGCTCCCAAGCAAGAACTATTTATGACATCAAGTTTA 1028
Qy      41 ProGlnAlaGlnLeuAlaSerIleTrypAsnLeuSerAspIleAspGlnAspGlyLysIeu 60
Db      1029 CCACAGGCTCAGCTGGCTTCATATGAAATCTTTCTGACATTCAGATGAGAAACTT 1088
Qy      61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db      1089 ACACAGAGGAATTTATCTTCGATGACACTCATTTATGATGCTATGTCTGGCCAAACA 1148
Qy      81 LeuProValLeuProProGluTyrIleProSerPheArgArg 96
Db      1149 CTGCCACCTGCTGCTGCCAGAAATACATTCACCTCTTTAGAGAA 1196

RESULT 11
AAZ39024
ID      AAZ39024 standard; cDNA; 5738 BP.
AC      AAZ39024;
DT      28-FEB-2000 (first entry)
XX      Mouse Esei1 cDNA sequence.
DE      Mouse Esei1 cDNA sequence.
KM      Mouse; murine; Esei1; Esei2; endocytosis; vesicular trafficking;
KW      regulation; actin cytoskeleton; detection; cancer; infection;
KM      EH-domain and SH3-domain regulator of endocytosis; anticancer;
KW      antiproliferative; antiviral; ss.
OS      Mus sp.
XX      WO9955728-A2.
PN      04-NOV-1999.
XX      27-APR-1999; 99WO-CA000375.
PF      27-APR-1998; 98CA-02230201.
PR      05-FEB-1999; 99US-0118739P.
XX      (HSCR-) HSC RES & DEV LP.
PI      Egan SE, Wang W, Sengar A;
XX      WPI: 2000-052802/04.
DR      P-PSDB; AA57449.
XX      New nucleic acid encoding Esei and 2 proteins, involved in regulation of
PT      endocytosis, used e.g. for treating cancer or preventing viral infection.
XX      Claim 6; Page 56-59; 99pp; English.
XX      The present invention specifically describes mammalian Esei and 2
CC      proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
CC      regulator of endocytosis). (I) are involved in regulation of clathrin-
CC      mediated endocytosis (as a complex with Esp15 protein), vesicular
CC      trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
CC      mimetics, fragments and inactive mutants); (I)-specific antibodies,
CC      sequences antisense to the (I) polynucleotide; agents that downregulate
CC      expression of Ese genes or antagonists of an Ese binding partner are used
CC      to treat diseases associated with undesirable endocytosis and resulting
CC      changes in cellular function. Particularly overexpression of Esei is used
CC      to block clathrin-mediated endocytosis in vivo or in cell cultures, while
CC      administration of (I) is used to promote endocytosis of selected cells.
CC      (Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
CC      cells that can be stimulated to proliferate by a growth factor receptor;
CC      and similar compounds (also inactive Ese mutants) can be used to prevent
CC      viral infection. Endocytosis may also be regulated, in vivo or in cell
```

```
CC      cultures, by forming an Ese-Esp15 complex, then binding dynamin to the
CC      complex. Generally conditions that can be treated include cancer;
CC      abnormal cell division or migration; viral infection; or abnormal
CC      receptor signalling, tissue development or synaptic transmission. The
CC      present sequence represents mouse Esei1 cDNA sequence
XX      Sequence 5738 BP; 1556 A; 1456 C; 1613 G; 1113 T; 0 U; 0 Other;
SQ      Alignment Scores:
Pred. No.: 5.3e-52 Length: 5738
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-720-934-2_copy_215_310 (1-96) x AAZ39024 (1-5738)
Qy      1 TrpAlaValProGlnSerSerArgLeuIleTyrArgGlnLeuPheAsnSerHisAspLys 20
Db      902 TGGGCTGTGCTCTAGTCAATCAAGCTGAATACAGGCAATTATTCATACAGCACACAAA 961
Qy      21 ThwterSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerIeu 40
Db      962 ACTATGAGTGACACTTACAGCTCCCAAGCAAGAACTATTTATGACATCAAGTTTA 1021
Qy      41 ProGlnAlaGlnLeuAlaSerIleTrypAsnLeuSerAspIleAspGlnAspGlyLysIeu 60
Db      1022 CCCACAGGCTCAGCTGGCTTCATATGAAATCTTTCTGACATTCAGATGAGAAACTC 1081
Qy      61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db      1082 ACTGAGAGGAATTTATCTTCGATGACACTCATTTATGATGCTATGTGCAATCTCGTACGCCA 1141
Qy      81 LeuProValLeuProProGluTyrIleProSerPheArgArg 96
Db      1142 CTGCCGCGCTGCTGCCAGAAATACATTCCTCTCTTACAGAA 1189

RESULT 12
AA180000
ID      AA180000 standard; cDNA; 548 BP.
AC      AA180000;
DT      06-NOV-2001 (first entry)
XX      Human polynucleotide SEQ ID NO 60.
DE      Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX      Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW      vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW      tissue growth factor; immunomodulatory; cancer; leukaemia;
KW      nervous system disorders; arthritis; inflammation; ss.
XX      Homo sapiens.
OS      Homo sapiens.
XX      WO200164835-A2.
PN      07-SEP-2001.
XX      26-FEB-2001; 2001WO-US004927.
PF      28-FEB-2000; 2000US-00515126.
PR      18-MAY-2000; 2000US-00577409.
XX      (HSE-) HSEQ INC.
PA      Tang YT, Liu C, Dermanac RT;
XX      WPI: 2001-514838/56.
DR      P-PSDB; AA000069.
XX      Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT      and treating e.g. leukemia, inflammation and immune disorders.
```

XX
PS Claim 1; SEQ ID NO 60; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 548 BP; 173 A; 124 C; 123 G; 127 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 9.77e-53 Length: 548
Score: 496.00 Matches: 95
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.00% Indels: 0
DB: 4 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x AA180000 (1-548)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
Db 53 TGGGCTTCTTCTCAGTATATACAGCTGAAATACAGGCAATATTATCACTAGCTATGACAAA 112

QY 21 ThMeSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerLeu 40
Db 113 ACTATGAGTGACACTTAACAGGCTCCCAAGCAAGAACTATTCTTATGAGTCAAGTTA 172

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
Db 173 CCACAGGCTCAGCTGGCTTCAATATGAAATCTTCTCAGCTTGTATCAAGATGAAAACTT 232

QY 61 ThrAlaGlnIlePheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 233 ACAGCAGAGCAATTATCTCGCAATGCACTCATTCATGATGCTAGTCTGGCCACCA 292

QY 81 LeuProValLeuLeuProGlnTyrIleProPheSerPheArg 95
Db 293 CTGCCACCTGTCTGCTCCAGATACATTCACCTCTTTTGA 337

RESULT 13
ABK43498
ID ABK43498 standard; cDNA; 3319 BP.
AC
XX
AC ABK43498;
XX
DT 05-JUN-2002 (first entry)
XX
DE DNA encoding novel central nervous system protein #78.
XX
XX Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy; gene; ss.
XX
XX Homo sapiens.
OS
XX
PN MO200155318-A2.

XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-06001332.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-020515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-021513P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234597P.
PR 25-SEP-2000; 2000US-0234598P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCT INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX

DR WPI; 2001-581633/65.
XX P-PSDB; JAU87166.
XX
XX New isolated nucleic acid encoding a protein for diagnosing, preventing,
PT treating or ameliorating medical conditions and used as food additives or
PT preservatives.
XX
XX Claim 1; SEQ ID NO 88; 837bp; English.
XX
XX The invention describes an isolated nucleic acid molecule (I) encoding a
CC novel central nervous system protein. (I) and polypeptides (III) encoded
CC by (I), are used to treat a medical conditions and in diagnosis of a
CC biological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC adenocarcinomas and irritable bowel syndrome, reproductive system
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,
Alignment Scores:
Pred. No.: 8.17e-51 Length: 3319
Score: 489.00 Matches: 96
Percent Similarity: 96.00% Conservative: 0
Best Local Similarity: 96.00% Mismatches: 0
Query Match: 97.60% Indels: 4
DB: Gaps: 1
US-09-720-934-2_COPY 215_310 (1-96) x ABK43498 (1-3319)
QY 1 TrrpAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
DB 801 TGGGCTTCTCCTAGTATCAAGACTGAATACAGGCATTATTATCATAGCATACAA 860
QY 21 ThrMetSerGlyHisLeuThr-----GlyProGlnAlaArgThrIleLeuMet 36
DB 861 ACTATGAGTGCACACTTAACAGGTTCTGTTAGGTCCTCCCAAGCAAGCAATTCATTATG 920
QY 37 GlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTyrPheLeuSerAspIleAspGln 56
DB 921 CAGTCAGCTTTACCAAGGCTCAGCTGCTTCATCAATATGAACTCTTCGACATTATCAA 980
QY 57 AspGlyLysLeuThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMet 76
DB 981 GATGGAAAACCTTAACGACGAGGAATTATCTCGGCAATGACACCTGATGATGACTATG 1040
QY 77 SerGlyGlnProLeuProProValLeuProProGlyTyrIleProProSerPheArgArg 96
DB 1041 TCTGGCCAAACCACTGCACCTGCTGCTCCAGAAATCAATTCACCTTTTGAAGA 1100
RESULT 14
AA163825
ID AA163825 standard; cDNA, 3466 BP.
XX
XX AA163825;
AC
XX
DT 22-OCT-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 33.
DE
XX

KW Human; antiarthritic; antirheumatic; antiproliferative; vasotropic;
KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; cytostatic; immunosuppressive; nootropic;
KW neuroprotective; antiallergic; hepatotropic; antidiabetic;
KW antiinflammatory; anticancer; vulnerrary; anticonvulsant; antibacterial;
KW antiparasitic; cardiant; gene therapy; cancer; immune disorder;
KW cardiovascular disorder; neurological disease; infection; human; ss.
XX
XX Homo sapiens.
XX
XX WO200155308-A2.
XX
XX
XX 02-AUG-2001.
XX
XX
XX 17-JAN-2001; 2001WO-US001309.
XX
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198133P.
XX 19-MAY-2000; 2000US-0205151P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 14-JUL-2000; 2000US-0218230P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226682P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 12-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0234984P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236337P.
PR 29-SEP-2000; 2000US-0236337P.
PR 29-SEP-2000; 2000US-0236337P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0251989P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM,
XX
XX WPI; 2001-488781/53.
XX P-PSDB; AAM43519.
XX
XX New isolated nucleic acids and polypeptides, useful for diagnosing,
XX treating and/or preventing human diseases and disorders.
XX
XX Claim 1; SEQ ID NO 33; 664bp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AA163803-AA164012) and
XX the encoded proteins (AAM434497-AAM43660) useful for preventing, treating
XX or ameliorating medical conditions e.g. by protein or gene therapy. The
XX genes were isolated from a range of human tissues disclosed in the
XX specification. The nucleic acids, proteins, antibodies and (ant)agonists
XX are useful in the diagnosis, treatment and prevention of: (a) cancer,
XX e.g. breast and ovarian cancer and other cancers of the adrenal gland,
XX b) immune disorders e.g. Addison's disease, allergies,
XX c) autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
XX Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d)
XX wound healing; (e) neurological diseases e.g. cerebral anoxia and
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
XX and parasitic infections. Note: The sequence data for this patent did not
XX form part of the printed specification, but was obtained in electronic
XX format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SO Sequence 3466 BP; 1056 A; 829 C; 861 G; 710 T; 0 U; 10 Other;
XX
XX Alignment Scores:
XX Pred. No.: 8.65e-51 Length: 3466
XX Score: 489.00 Matches: 96
XX Percent Similarity: 96.00% Conservative: 0
XX Best Local Similarity: 96.00% Mismatches: 0
XX Query Match: 97.60% Indels: 4
XX DB: Gaps: 1
XX
XX US-09-720-934-2_COPY_215_310 (1-96) x AA163825 (1-3466)
XX
XX QY 1 TTPALAVAlProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
XX Db 919 TGGGCTGTTCCATCATCAAGACTGAAATPACAGGCAATTTATCAATGATCATACAAA 978
XX QY 21 ThnMetSerGlyHisLeuThr-----GlyProGlnAlaArgThrIleLeuMet 36
XX Db 979 ACTATGAGTGGACACTTAACAGGTTCCTTAGGTCCCAAGCAAGAACTATTCTTAG 1038
XX QY 37 GlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTyrPheAsnSerHisAspLys 56
XX Db 1039 CAGTCMAATTACCAACAGGCTAGCTGGCTTCMAATGAAATCTTTCGACATTGATCA 1098
XX QY 57 AspGlyLysLeuThrAlaGlnIlePheIleLeuAlaMetHisLeuIleAspValAlaMet 76
XX Db 1099 GATGGAAAACCTTACCGACAGAGAAATTTATCTCGGCAATACACCTCATGATTAAGTANG 1158
XX QY 77 SerGlyGlnProLeuProProValLeuProProGlnTyrIleProProSerPheArgArg 96
XX Db 1159 TCTGGCCAAACCACTCCCACTTCTCGCTCCAGATATACATTCACCTCTTTAGAGA 1218
XX
XX RESULT 15
XX AAS84763
XX ID AAS84763 standard; cDNA; 7435 BP.
XX

AC AAS84763;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #20567.
XX
XX DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US008631.
XX
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX P-PSDB; ABG20576.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 1; SEQ ID NO 20567; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food imaging
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 7435 BP; 2190 A; 1690 C; 1829 G; 1726 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 3.44e-41 Length: 7435
XX Score: 416.00 Matches: 96
XX Percent Similarity: 93.20% Conservative: 0
XX Best Local Similarity: 93.20% Mismatches: 0
XX Query Match: 83.03% Indels: 7
XX DB: Gaps: 0
XX
XX US-09-720-934-2_COPY_215_310 (1-96) x AAS84763 (1-7435)
XX
XX QY 1 TTPALAVAlProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAsp 19
XX Db 907 TGGGCTGTTCCATCATCAAGACTGAAATPACAGGCAATTTATTCATATGATCATGAC 966
XX QY 20 LysThr--MetSerGlyHis-LeuThr-GlyProGlnAlaArgThrIleLeuMetClnse 38
XX

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:03:00 ; Search time 67.2225 Seconds
(without alignments)
792,522 Million cell updates/sec

Title: US-09-720-934-2_COPY_215_310
Perfect score: 501
Sequence: 1 WAPVQSSRLKRLQLENSHDK.....SGQPLPPVLPPEYTPSPRR 96

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frim+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09720934/runat_30062004_064540_1345/app_query.fasta_1.1386
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPT=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTMT=plo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USPR=US09720934 @CGN 1 1 284 @runat_30062004_064540_13455 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A COMB.seg:*
2: /cgn2_6/prodata/2/ina/5B COMB.seg:*
3: /cgn2_6/prodata/2/ina/6A COMB.seg:*
4: /cgn2_6/prodata/2/ina/6B COMB.seg:*
5: /cgn2_6/prodata/2/ina/FACTUS COMB.seg:*
6: /cgn2_6/prodata/2/ina/backfile1.seg:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	349	69.7	2017	4	US-09-404-879A-72
2	349	69.7	2017	4	US-09-358-933-72
3	349	69.7	2017	4	US-09-215-681-72
4	349	69.7	2017	4	US-09-216-003A-72
5	208	41.5	480	4	US-09-404-879A-60
6	208	41.5	480	4	US-09-338-933-60
7	208	41.5	480	4	US-09-215-681-60
8	208	41.5	480	4	US-09-216-003A-60
9	208	41.5	531	4	US-09-404-879A-5
10	208	41.5	531	4	US-09-338-933-5
11	208	41.5	531	4	US-09-215-681-5
12	208	41.5	531	4	US-09-216-003A-5

13	204.5	40.8	3033	1	US-08-095-737-3	Sequence 3, Appl1
14	204.5	40.8	3033	2	US-08-480-145-3	Sequence 3, Appl1
15	204.5	40.8	3033	1	US-08-477-389-3	Sequence 3, Appl1
16	197.5	39.4	4165	1	US-08-095-737-1	Sequence 1, Appl1
17	197.5	39.4	4165	1	US-08-480-145-1	Sequence 1, Appl1
18	197.5	39.4	4165	2	US-08-477-389-1	Sequence 2, Appl1
19	189	37.7	3348	2	US-09-312-762A-2	Sequence 2, Appl1
20	189	37.7	3508	4	US-09-312-762A-1	Sequence 1, Appl1
21	189	37.7	14707	4	US-09-312-762A-3	Sequence 3, Appl1
22	176	35.1	3635	4	US-09-312-762A-6	Sequence 6, Appl1
23	170	33.9	2373	4	US-09-023-655-850	Sequence 850, App
24	162	32.3	590	3	US-09-328-111-211	Sequence 211, App
25	151	30.1	2815	4	US-09-312-762A-7	Sequence 7, Appl1
26	138	27.5	5764	4	US-09-312-762A-7	Sequence 8, Appl1
27	94.5	18.9	845	3	US-08-998-416-543	Sequence 543, App
28	94	18.8	503	4	US-09-370-838-148	Sequence 1048, App
29	90.5	18.1	884	4	US-09-833-381-1048	Sequence 1048, App
30	87	17.4	853	3	US-08-963-409-2	Sequence 2, Appl1
31	81.5	16.3	1776	3	US-08-655-352-10	Sequence 10, Appl1
32	81.5	16.3	1776	4	US-09-258-016-10	Sequence 10, Appl1
33	81.5	16.3	1776	4	US-09-257-825B-10	Sequence 8, Appl1
34	81	16.2	1201	3	US-09-048-889-8	Sequence 8, Appl1
35	81	16.2	2259	4	US-09-399-913-71	Sequence 69, Appl1
36	81	16.2	2263	4	US-09-399-913-69	Sequence 48, Appl1
37	81	16.2	2413	4	US-09-399-913-48	Sequence 50, Appl1
38	79	15.8	1835	4	US-09-338-671-3	Sequence 31, Appl1
39	79	15.8	1835	3	US-09-399-913-31	Sequence 35, Appl1
40	78	15.6	2841	4	US-09-399-913-35	Sequence 35, Appl1
41	76	15.2	2644	4	US-09-298-731-35	Sequence 5, Appl1
42	76	15.2	2644	2	US-08-818-252-5	Sequence 5, Appl1
43	75	15.0	1929	3	US-08-818-252-5	Sequence 7, Appl1
44	75	15.0	1929	2	US-08-818-253-7	Sequence 7, Appl1
45	75	15.0	1971	2	US-08-818-253-7	Sequence 7, Appl1

ALIGNMENTS

RESULT 1
US-09-404-879A-72
; Sequence 72, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404, 879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-404-879A-72

Alignment Scores:
Pred. No.: 1,35e-37
Score: 349.00
Percent Similarity: 82.11%
Best Local Similarity: 71.58%
Query Match: 69.66%
DB: 4
Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-09-404-879A-72 (1-2017)
QY 1 TYPAlaValProGlnSerSerArgLeuLyTYArgGlnLeuPheAnsSerHisAspLys 20
DB 753 TGGGCGAGTTCCTCAGCGCTTCAAGATTAAAGTATACGGCAAAATTTATACCTCAGACAAA 812
QY 21 ThMeSerSelYHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40

```
Db 813 GGCATGACGGATACCTCTCAGCTTTCAAGCTAATAATGCCCTTCTTCAGTCAAAATCTC 872
Qy 41 ProglinalaginLeuA1aSer11eTrrasnLeuSerAap11eAspGlnAspGlyLysLeu 60
Db 873 TCTCAAACTCAGCTAGCTACTATTGAGCTCTGGCTGACATCGATGAGTGGACGACAGTTG 932
Qy 61 ThrAlaglugluPhe11eLeuA1aMeth1sLeu11eAspVal1a1aMetSerGlyGlnPro 80
Db 933 AAAGCTGAAGAATTATTCTTGGCGATGACCTCACTGACATGCGCAAAAGCTGGACAGCCA 992
Qy 81 LeuProProVal11eLeuProProGluTyr11eProProSerPheArg 95
Db 993 CTACCACTGACGTTGCTCCCGAGCTGTGCTCCCTCATCTTTTCAGA 1037

RESULT 2
US-09-338-933-72
; Sequence 72, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338,933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-72

Alignment Scores:
Pred. No.: 1,35e-37 Length: 2017
Score: 349.00 Matches: 68
Percent Similarity: 82.11% Conservative: 10
Best Local Similarity: 71.58% Mismatches: 17
Query Match: 69.66% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-09-338-933-72 (1-2017)
Qy 1 TrrAlava1ProglinsSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
Db 753 TGGGAGATTCTCAGCTTCAAGATTAAAGTATCGGCAAAAATTAAATGATCTAGACAAA 812
Qy 21 ThrMetSerGlyHisLeuThrg1yProGln1a1aArgThr11eLeuMetGlnSerSerLeu 40
Db 813 GGCATGACGGATACCTCTCAGCTTTCAAGCTAATAATGCCCTTCTTCAGTCAAAATCTC 872
Qy 41 ProglinalaginLeuA1aSer11eTrrasnLeuSerAap11eAspGlnAspGlyLysLeu 60
Db 873 TCTCAAACTCAGCTAGCTACTATTGAGCTCTGGCTGACATCGATGAGTGGACGACAGTTG 932
Qy 61 ThrAlaglugluPhe11eLeuA1aMeth1sLeu11eAspVal1a1aMetSerGlyGlnPro 80
Db 933 AAAGCTGAAGAATTATTCTTGGCGATGACCTCACTGACATGCGCAAAAGCTGGACAGCCA 992
Qy 81 LeuProProVal11eLeuProProGluTyr11eProProSerPheArg 95
Db 993 CTACCACTGACGTTGCTCCCGAGCTGTGCTCCCTCATCTTTTCAGA 1037

RESULT 3
US-09-215-681-72
; Sequence 72, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Fridakis, Tony N.
; APPLICANT: King, Gordon E.
```

```
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; FILE REFERENCE: 210121.463
; CURRENT APPLICATION NUMBER: US/09/215,681A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-215-681-72

Alignment Scores:
Pred. No.: 1,35e-37 Length: 2017
Score: 349.00 Matches: 68
Percent Similarity: 82.11% Conservative: 10
Best Local Similarity: 71.58% Mismatches: 17
Query Match: 69.66% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-09-215-681-72 (1-2017)
Qy 1 TrrAlava1ProglinsSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
Db 753 TGGGAGATTCTCAGCTTCAAGATTAAAGTATCGGCAAAAATTAAATGATCTAGACAAA 812
Qy 21 ThrMetSerGlyHisLeuThrg1yProGln1a1aArgThr11eLeuMetGlnSerSerLeu 40
Db 813 GGCATGACGGATACCTCTCAGCTTTCAAGCTAATAATGCCCTTCTTCAGTCAAAATCTC 872
Qy 41 ProglinalaginLeuA1aSer11eTrrasnLeuSerAap11eAspGlnAspGlyLysLeu 60
Db 873 TCTCAAACTCAGCTAGCTACTATTGAGCTCTGGCTGACATCGATGAGTGGACGACAGTTG 932
Qy 61 ThrAlaglugluPhe11eLeuA1aMeth1sLeu11eAspVal1a1aMetSerGlyGlnPro 80
Db 933 AAAGCTGAAGAATTATTCTTGGCGATGACCTCACTGACATGCGCAAAAGCTGGACAGCCA 992
Qy 81 LeuProProVal11eLeuProProGluTyr11eProProSerPheArg 95
Db 993 CTACCACTGACGTTGCTCCCGAGCTGTGCTCCCTCATCTTTTCAGA 1037

RESULT 4
US-09-216-003A-72
; Sequence 72, Application US/09216003A
; Patent No. 6670463
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Fridakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
; FILE REFERENCE: 210121.462
; CURRENT APPLICATION NUMBER: US/09/216,003A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-216-003A-72

Alignment Scores:
Pred. No.: 1,35e-37 Length: 2017
Score: 349.00 Matches: 68
Percent Similarity: 82.11% Conservative: 10
Best Local Similarity: 71.58% Mismatches: 17
Query Match: 69.66% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-09-216-003A-72 (1-2017)
```


STATE: California
COUNTRY: United States of America
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/095,737
FILING DATE: 19930722
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned A
REGISTRATION NUMBER: 29,655
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-0176
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3033 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 111..2802
US-08-095-737-3

Alignment Scores:
Pred. NO.: 1.14e-17 Length: 3033
Score: 204.50 Matches: 40
Percent Similarity: 60.42% Conservative: 18
Best Local Similarity: 41.67% Mismatches: 37
Query Match: 40.82% Indels: 1
DB: 1 Gaps: 1

US-09-720-934-2_COPY_215_310 (1-96) x US-08-095-737-3 (1-3033)
QY 1 ThrAlaValProGlnSerSerArgLeuIysTyrArgGlnLeuPheAsnSerHisAspLys 20
DB 474 TGGGCTGTAAGTCTGAAGATTAAGCCAAATATGATGCATTTTGGACAGTTTA---AGC 530
QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
DB 531 CCAGTGGATGATTTTGTCTGCGATTAAGTGAACACAGTGTGCTCACTTAAGTTA 590
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysIeu 60
DB 591 CCTGTGGAATCCTTGGAAAGATTGGAGATTGATGATTATTTGACACAGATGGAAGAGCTG 650
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 651 GACAGAGATGAGTTTGCAGTTGCTGTTTGGATATCTGTGACCTGGAAGAAAGAACTT 710
QY 81 LeuProValLeuProGluTyrIleProProSerPheArg 96
DB 711 GTGCCAATGTCCTTGCCTCCAGCCTTGTGTGCACCTTAAAGAGAAA 758

RESULT 14
US-08-480-145-3
; Sequence 3, Application US/08480145
; Patent No. 5717067
; GENERAL INFORMATION:
; APPLICANT: DiFiore, Pier P
; APPLICANT: Razioli, Francesca
; TITLE OF INVENTION: A Substrate for the Epidermal Growth
; FACTOR RECEPTOR KINASE
; NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: California
COUNTRY: United States of America
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,145
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/095,737
FILING DATE: 22-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH060.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-0176
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3033 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 111..2802
US-08-480-145-3

Alignment Scores:
Pred. NO.: 1.14e-17 Length: 3033
Score: 204.50 Matches: 40
Percent Similarity: 60.42% Conservative: 18
Best Local Similarity: 41.67% Mismatches: 37
Query Match: 40.82% Indels: 1
DB: 1 Gaps: 1

US-09-720-934-2_COPY_215_310 (1-96) x US-08-480-145-3 (1-3033)
QY 1 ThrAlaValProGlnSerSerArgLeuIysTyrArgGlnLeuPheAsnSerHisAspLys 20
DB 474 TGGGCTGTAAGTCTGAAGATTAAGCCAAATATGATGCATTTTGGACAGTTTA---AGC 530
QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
DB 531 CCAGTGGATGATTTTGTCTGCGATTAAGTGAACACAGTGTGCTCACTTAAGTTA 590
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysIeu 60
DB 591 CCTGTGGAATCCTTGGAAAGATTGGAGATTGATGATTATTTGACACAGATGGAAGAGCTG 650
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 651 GACAGAGATGAGTTTGCAGTTGCTGTTTGGATATCTGTGACCTGGAAGAAAGAACTT 710
QY 81 LeuProValLeuProGluTyrIleProProSerPheArg 96
DB 711 GTGCCAATGTCCTTGCCTCCAGCCTTGTGTGCACCTTAAAGAGAAA 758

RESULT 15
US-08-477-389-3
; Sequence 3, Application US/08477389

Patent No. 5872219
GENERAL INFORMATION:
APPLICANT: DiFiore, Pier P
APPLICANT: Fazioli, Francesca
TITLE OF INVENTION: A Substrate for the Epidermal Growth
NUMBER OF INVENTION: Factor Receptor Kinase
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: California
COUNTRY: United States of America
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,389
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/095,737
FILING DATE: 22-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH060.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3033 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 111..2802
US-08-477-389-3

Alignment Scores:
Pred. No.: 1,14e-17 Length: 3033
Score: 204.50 Matches: 40
Percent Similarity: 60.42% Conservative: 18
Best Local Similarity: 41.67% Mismatches: 37
Query Match: 40.82% Indels: 1
DB: 2 Gaps: 1

US-09-720-934-2_COPY_215_310 (1-96) x US-08-477-389-3 (1-3033)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTYrArgGlnLeuPheAsnSerHisAspLys 20
Db 474 TGGGCTGTAAAGCTGAAGTAAAGCCAAATATGATGCAATTTTGAACAGTTTA--AGC 530
QY 21 ThMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 531 CCAGTGAATGATTTTGTCTGTGTGATTAAGTGAACCAAGTGTGCTCAACTCTAAGTTA 590
QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
Db 591 CCTGTGGAATCTCTTGAAGAGTTTGGAGTTGAGTGATTTGACCAAGATGAAAGCTG 650
QY 61 ThrAlaGlnGlnPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 651 GACAGAGATGAGTTTGCAGTGTCCATGTTTGTGTACTGTGCACTGAGAAAGAACT 710

QY 81 LeuProProValLeuProProGlnUtyrIleProProSerPheArgArg 96
Db 711 GTGCCAATGTCTTGTGCTCCAGGCTTGTGTGCCACCTTTTAAGAGAAA 758
Search completed: July 1, 2004, 19:53:41
Job time : 71.2225 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:09:20 ; Search time 294.52 Seconds
(without alignments)
1572.471 Million cell updates/sec

Title: US-09-720-934-2_COPY_215_310
Perfect score: 501
Sequence: 1 MAVPQSSRLKXRLPNSHDK.....SQGPLPVLPPEITPSPFR 96

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3163042 seqs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O/cgn2_1/USPTO.spool_p/US09720934/runat_30062004_064541_13481/app_query_faasta_1.1386
-DB=Published Applications NA -OPMT=fastrep -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09720934.@CGN_1_1500.@runat_30062004_064541_13481
-NCPU=6 -ICPU=3 -NO MAP -LARGEOUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :
Published Applications NA:
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result Query SUMMARIES
No. Score Match Length DB ID Description

1	489	97.6	3319	11	US-09-764-875-88	Sequence 88, Appl
2	489	97.6	3466	16	US-10-158-057-33	Sequence 33, Appl
3	349	69.7	2017	9	US-09-884-441-72	Sequence 72, Appl
4	349	69.7	2017	10	US-09-907-969-72	Sequence 72, Appl
5	349	69.7	2017	10	US-09-827-271-72	Sequence 72, Appl
6	349	69.7	2017	15	US-10-198-053-72	Sequence 72, Appl
7	345	68.9	5828	13	US-10-398-885A-15	Sequence 15, Appl
8	334	66.7	955	10	US-09-764-881-54	Sequence 54, Appl
9	334	66.7	955	11	US-09-764-881-54	Sequence 54, Appl
10	334	66.7	955	13	US-09-764-881-54	Sequence 54, Appl
11	334	66.7	955	16	US-10-242-747-54	Sequence 54, Appl
12	334	66.7	955	16	US-10-158-057-33	Sequence 34, Appl
13	218.5	43.6	3976	16	US-10-369-493-36727	Sequence 36727, A
14	208	41.5	480	9	US-09-884-441-60	Sequence 60, Appl
15	208	41.5	480	10	US-09-907-969-60	Sequence 60, Appl
16	208	41.5	480	10	US-09-827-271-60	Sequence 60, Appl
17	208	41.5	480	15	US-10-198-053-60	Sequence 60, Appl
18	208	41.5	531	9	US-09-884-441-5	Sequence 5, Appl
19	208	41.5	531	10	US-09-907-969-5	Sequence 5, Appl
20	208	41.5	531	10	US-09-827-271-5	Sequence 5, Appl
21	208	41.5	531	15	US-10-198-053-5	Sequence 5, Appl
22	191	38.1	2259	16	US-10-369-493-26020	Sequence 26020, A
23	189	37.7	462	10	US-09-918-995-21728	Sequence 21728, A
24	189	37.7	605	12	US-10-152-319A-1126	Sequence 1126, Ap
25	189	37.7	3348	9	US-09-312-762A-2	Sequence 2, Appl
26	189	37.7	3508	9	US-09-312-762A-1	Sequence 1, Appl
27	189	37.7	14707	9	US-09-312-762A-3	Sequence 3, Appl
28	188	37.5	860	13	US-10-027-633-165796	Sequence 165796,
29	188	37.5	860	13	US-10-027-633-165797	Sequence 165797,
30	188	37.5	860	16	US-10-027-633-165796	Sequence 165796,
31	188	37.5	860	16	US-10-027-633-165797	Sequence 165797,
32	188	37.5	3485	15	US-10-171-581-354	Sequence 354, App
33	187.5	37.4	2442	9	US-09-964-899-24	Sequence 24, Appl
34	187	37.3	662	13	US-10-027-633-290824	Sequence 290824,
35	187	37.3	523	10	US-10-027-633-290824	Sequence 290824,
36	185	36.9	523	10	US-09-814-353-14688	Sequence 14688, A
37	185	36.9	707	10	US-09-764-881-186	Sequence 186, App
38	185	36.9	707	13	US-09-764-881-186	Sequence 186, App
39	185	36.9	707	16	US-10-242-747-186	Sequence 186, App
40	185	36.9	707	16	US-10-158-057-373	Sequence 373, App
41	185	36.9	2105	11	US-09-764-875-175	Sequence 175, App
42	185	36.9	2414	10	US-09-764-875-20215	Sequence 20215, A
43	185	36.9	2907	16	US-10-108-260A-1761	Sequence 1761, Ap
44	185	36.9	3017	9	US-09-818-143-8	Sequence 8, Appl
45	184	36.7	523	10	US-09-814-353-1957	Sequence 1957, Ap

ALIGNMENTS

RESULT 1
US-09-764-875-88
; Sequence 88, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0202
; CURRENT APPLICATION NUMBER: US/09/764,875
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 88
; LENGTH: 3319
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-875-88

Alignment Scores:
Pred. No.: 8.22e-60
Score: 489.00
Percent Similarity: 96.00%

Length: 3319
Matches: 96
Conservative: 0

Best Local Similarity: 96.00%
Query Match: 97.60%
DB: 11
Gaps: 1

US-09-720-934-2_COPY_215_310 (1-96) x US-09-764-875-88 (1-3319)

QY 1 TrpAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
DB 801 TGGGCTGTTCCTCACTCATCAAGACTGAATAATACAGGCAATTATTAATAGTCATACAA 860

QY 21 ThrMetSerGlyHisLeuThr-----GlyProGlnAlaArgThrIleLeuMet 36
DB 861 ACTAGAGCTGACACCTTAACAGCTTCTCTGTAGGCTCCCAAGCAACTATTCTTATG 920

QY 37 GlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTyrPheLeuSerAspIleAspGln 56
DB 921 CAGTCAGATTACACAGAGCTGCTGCTGCTCAATATGAAATCTTTCAGCACTTGATCAA 980

QY 57 AspGlyLysLeuThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMet 76
DB 981 GATGGAAACTTACAGCAGAGGATTTATCTCGGCAATGCACCTCATTTGATGACTATG 1040

QY 77 SerGlyGlnProLeuProProValLeuProProGluTyrIleProProSerPheArgArg 96
DB 1041 TCTGGCCACCACTGCGCACTGCTCTGCTCCAGAAATCATTCACCTTTTGAAGA 1100

RESULT 2
US-10-158-057-33
; Sequence 33, Application US/10158057
; Publication No. US20040014039A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1205C1
; CURRENT APPLICATION NUMBER: US/10/158,057
; CURRENT FILING DATE: 2002-06-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33
; LENGTH: 3466
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3194)
; OTHER INFORMATION: n equals a,t,b, or c
; NAME/KEY: misc feature
; LOCATION: (3465)
; OTHER INFORMATION: n equals a,t,b, or c
US-10-158-057-33

Alignment Scores:
Pred. No.: 8.73e-60 Length: 3466
Score: 489.00 Matches: 96
Percent Similarity: 96.00% Conservative: 0
Best Local Similarity: 96.00% Mismatches: 0
Query Match: 97.60% Indels: 4
DB: 16 Gaps: 1

US-09-720-934-2_COPY_215_310 (1-96) x US-10-158-057-33 (1-3466)

QY 1 TrpAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
DB 919 TGGGCTGTTCCTCACTCATCAAGACTGAATAATACAGGCAATTATTAATAGTCATACAA 978

QY 21 ThrMetSerGlyHisLeuThr-----GlyProGlnAlaArgThrIleLeuMet 36
DB 979 ACTATAGTGGACACTTAACAGCTTCTGTAGTCCCAAGCAAGACTATTCTTATG 1038

QY 37 GlnSerSerLeuProGlnAlaGlnLeuAlaSerIleTyrPheLeuSerAspIleAspGln 56

DB 1039 CAGTCAAGTTTACACAGAGCTGAGCTTCATATGGAATCTTTCGACATTGATCAA 1096

QY 57 AspGlyLysLeuThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMet 76
DB 1099 GATGGAAACTTACAGCAGAGGAAATTTATCTCGGCAATGCACCTCATTTGATGACTATG 1158

QY 77 SerGlyGlnProLeuProProValLeuProProGluTyrIleProProSerPheArgArg 96
DB 1159 TCTGGCCACCACTGCGCACTGCTCTGCTCCAGAAATCATTCACCTTTTGAAGA 1218

RESULT 3
US-09-884-441-72
; Sequence 72, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Carter, Darick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C7
; CURRENT APPLICATION NUMBER: US/09/884,441
; CURRENT FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-884-441-72

Alignment Scores:
Pred. No.: 7.22e-40 Length: 2017
Score: 349.00 Matches: 68
Percent Similarity: 82.11% Conservative: 10
Best Local Similarity: 71.58% Mismatches: 17
Query Match: 69.66% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-09-884-441-72 (1-2017)

QY 1 TrpAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
DB 753 TGGGCACTTCTCACTCACTTCAAGATTAAGATGCGCAAAATTTAATAGTCTAGACAA 812

QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
DB 813 GGCATGAGCGGATACCTCTCAAGGTTTCAAGCTAGAAATGCCCTTTCAGTCAAAATCTC 872

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPheAsnSerAspIleAspGlnAspGlyLysLeu 60
DB 873 TCTGAACCTCAGCTAGCTACATTTTGGACTGTGCTGACATCGATGAGACGACAGTTG 932

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
DB 933 AAAGCTAAGAAATTTATCTGCGCATGACACTCACTAAGGCGCAAGCTGACAGCA 992

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArg 95
DB 993 CTACCACTGACGTTGCTCCGAGCTTGTCTCTCAATCTTTCAGA 1037

RESULT 4
US-09-907-969-72
; Sequence 72, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Ranger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.

```
; APPLICANT: Carter, Darrick
; APPLICANT: Hill, Paul
; APPLICANT: Albone, Earl
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/09/907,969
; CURRENT FILING DATE: 2001-07-17
; NUMBER OF SEQ ID NOS: 596
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-907-969-72

Alignment Scores:
Pred. No.: 7.22e-40 Length: 2017
Score: 349.00 Matches: 68
Percent Similarity: 82.11% Conservative: 10
Best Local Similarity: 71.58% Mismatches: 17
Query Match: 69.66% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-09-907-969-72 (1-2017)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
   |||||
Db 753 TGGGAGGTTCCCTCAGCTTCAGATTAAAGTATCGGCAAAATTTATATAGCTCAGACAA 812

QY 21 ThrMetSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
   |||||
Db 813 GGCATAGCGGAGATACCTCTCAGGTTTCAAGCTAGAAATGCCCTCTTCAGTCAAAATCTC 872

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPasnLeuSerAspIleAspGlnAspGlyLysLeu 60
   |||||
Db 873 TCTCAAACTCAGCTAGTACTATTGGACTCTGGCTGACATCGATGTCAGACAGCTTG 932

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
   |||||
Db 933 AAAGCTGAAGAATTATTTCTGGCGATGCACCTCAGCTGACATGGCCAAAGCTGGAGCACA 992

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArg 95
   |||||
Db 993 CTACCACTGACGTTGCTCCGAGCTTTCCTTCATCTTCACGA 1037

RESULT 5
US-09-827-271-72
; Sequence 72, Application US/09827271
; Publication No. US20030165504A1
; GENERAL INFORMATION:
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C6
; CURRENT APPLICATION NUMBER: US/09/827,271
; CURRENT FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-827-271-72

Alignment Scores:
Pred. No.: 7.22e-40 Length: 2017
Score: 349.00 Matches: 68
Percent Similarity: 82.11% Conservative: 10
Best Local Similarity: 71.58% Mismatches: 17
Query Match: 69.66% Indels: 0
DB: 10 Gaps: 0
```

```
US-09-720-934-2_COPY_215_310 (1-96) x US-09-827-271-72 (1-2017)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
   |||||
Db 753 TGGGAGGTTCCCTCAGCTTCAGATTAAAGTATCGGCAAAATTTATATAGCTCAGACAA 812

QY 21 ThrMetSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
   |||||
Db 813 GGCATAGCGGAGATACCTCTCAGGTTTCAAGCTAGAAATGCCCTCTTCAGTCAAAATCTC 872

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPasnLeuSerAspIleAspGlnAspGlyLysLeu 60
   |||||
Db 873 TCTCAAACTCAGCTAGTACTATTGGACTCTGGCTGACATCGATGTCAGACAGCTTG 932

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
   |||||
Db 933 AAAGCTGAAGAATTATTTCTGGCGATGCACCTCAGCTGACATGGCCAAAGCTGGAGCACA 992

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArg 95
   |||||
Db 993 CTACCACTGACGTTGCTCCGAGCTTTCCTTCATCTTCACGA 1037

RESULT 6
US-10-198-053-72
; Sequence 72, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198,053
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-198-053-72

Alignment Scores:
Pred. No.: 7.22e-40 Length: 2017
Score: 349.00 Matches: 68
Percent Similarity: 82.11% Conservative: 10
Best Local Similarity: 71.58% Mismatches: 17
Query Match: 69.66% Indels: 0
DB: 15 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-10-198-053-72 (1-2017)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
   |||||
Db 753 TGGGAGGTTCCCTCAGCTTCAGATTAAAGTATCGGCAAAATTTATATAGCTCAGACAA 812

QY 21 ThrMetSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
   |||||
Db 813 GGCATAGCGGAGATACCTCTCAGGTTTCAAGCTAGAAATGCCCTCTTCAGTCAAAATCTC 872

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPasnLeuSerAspIleAspGlnAspGlyLysLeu 60
   |||||
Db 873 TCTCAAACTCAGCTAGTACTATTGGACTCTGGCTGACATCGATGTCAGACAGCTTG 932

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
   |||||
Db 933 AAAGCTGAAGAATTATTTCTGGCGATGCACCTCAGCTGACATGGCCAAAGCTGGAGCACA 992

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArg 95
   |||||
Db 993 CTACCACTGACGTTGCTCCGAGCTTTCCTTCATCTTCACGA 1037
```

Db 993 CTACCACTGACGTTGCCCTCCGAGCTGTCTCCCTCATCTTTACAGA 1037

RESULT 7

US-10-398-885A-15

Sequence 15, Application US/1039888B5A

Publication No. US20040053282A1

GENERAL INFORMATION:

APPLICANT: Sugita, Yuji

APPLICANT: Hashida, Ryoichi

APPLICANT: Ogawa, Kaoru

APPLICANT: Nagasu, Takeshi

APPLICANT: Obayashi, Masaya

APPLICANT: Saito, Hirohisa

APPLICANT: Takahashi, Eiki

TITLE OF INVENTION: Method of Testing For Allergic Diseases

FILE REFERENCE: SHIMIZU-07907

CURRENT APPLICATION NUMBER: US/10/398, 885A

CURRENT FILING DATE: 2003-08-11

PRIOR APPLICATION NUMBER: PCT/JP01/08937

PRIOR FILING DATE: 2001-10-11

PRIOR APPLICATION NUMBER: JP 2000-314093

PRIOR FILING DATE: 2000-10-13

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn version 3.1

SEQ ID NO 15

LENGTH: 5828

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthetic

FEATURE:

NAME/KEY: CDS

LOCATION: (7)..(5052)

OTHER INFORMATION:

US-10-398-885A-15

Alignment Scores:

Pred. No.: 1.19e-38 Length: 5828

Score: 345.00 Matches: 65

Percent Similarity: 83.16% Conservative: 14

Best Local Similarity: 68.42% Mismatches: 16

Query Match: 68.86% Indels: 0

Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-10-398-885A-15 (1-5828)

QY 1 TrrAlaValProGlnSerSerArgLeuIysTyrArgGlnLeuPheAsnSerHisAspIys 20

Db 751 TGGGACAGTCTCTCAGCCTACAAAGATTAAATATCGGCAAAATTTAATACCTTGACAAA 810

QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40

Db 811 AGTATGATGTGATATCTCTCAGGTTTTCAGCTAGAAATGCCCTTCTTCAGTCAAAATCTT 870

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysIeu 60

Db 871 TCTCAACTCAGCTGGCTGACTACTTGGACTCTGGCTGACCTTGATGATGATGACAGCTA 930

QY 61 ThrAlaGlnGlnPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80

Db 931 AAACGAGAGAGGTTATTCTTGCAATGACCTTACTGACATGCGCCAAAGCTGACAGCCA 990

QY 81 LeuProProValLeuProProGlnTyrIleProProSerPheArg 95

Db 991 TTACCACTGACTTACCTCCGAGCTGTCTCCCTCATCTTTACAG 1035

RESULT 8

US-09-764-881-54

Sequence 54, Application US/097648B1

Publication No. US20030125246A9

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PT207

CURRENT APPLICATION NUMBER: US/09/764, 881

CURRENT FILING DATE: 2001-01-17

Prior application data removed - refer to PALM or file wrapper

NUMBER OF SEQ ID NOS: 192

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 54

LENGTH: 955

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (550)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE

LOCATION: (934)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE

LOCATION: (942)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE

LOCATION: (944)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE

LOCATION: (948)

OTHER INFORMATION: n equals a,t,g, or c

US-09-764-881-54

Alignment Scores:

Pred. No.: 3.78e-38 Length: 955

Score: 334.00 Matches: 64

Percent Similarity: 81.05% Conservative: 13

Best Local Similarity: 67.37% Mismatches: 18

Query Match: 66.67% Indels: 0

Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x US-09-764-881-54 (1-955)

QY 1 TrrAlaValProGlnSerSerArgLeuIysTyrArgGlnLeuPheAsnSerHisAspIys 20

Db 390 TGGGACAGTCTCTCAGCCTACAAAGATTAAATATCGGCAAAATTTAATACCTTGACAAA 449

QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40

Db 450 AGTATGATGTGATATCTCTCAGGTTTTCAGCTAGAAATGCCCTTCTTCAGTCAAAATCTT 509

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysIeu 60

Db 510 TCTCAACTCAGCTGGCTGACTACTTGGACTCTGGCTGACCTTGATGATGATGACAGCTA 569

QY 61 ThrAlaGlnGlnPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80

Db 570 AAACGAGAGAGGTTATTCTTGCAATGACCTTACTGACATGCGCCAAAGCTGACAGCCA 629

QY 81 LeuProProValLeuProProGlnTyrIleProProSerPheArg 95

Db 630 TTACCACTGACTTACCTCCGAGCTGTCTCCCTCATCTTTACAG 674

RESULT 9

US-09-764-875-91

Sequence 91, Application US/09764875

Publication No. US20040018969A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: P0202

CURRENT APPLICATION NUMBER: US/09/764, 875

CURRENT FILING DATE: 2001-01-17

Prior application data removed - consult PALM or file wrapper

NUMBER OF SEQ ID NOS: 1249

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 91

LENGTH: 955


```
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (550)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (934)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (942)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (944)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (948)
/ OTHER INFORMATION: n equals a,t,g, or c
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-764-875-91
```

```
Alignment Scores:
Pred. No.: 3.78e-38 Length: 955
Score: 334.00 Matches: 64
Percent Similarity: 81.05% Conservative: 13
Best Local Similarity: 67.37% Mismatches: 18
Query Match: 66.67% Indels: 0
DB: 11 Gaps: 0
```

US-09-720-934-2_COPY_215_310 (1-96) x US-09-764-875-91 (1-955)

```
QY 1 TrrAlvAlProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
Db 390 TGGGCGCTTCTAGCCTACAGATTAAATATCGCAAAATTTAACTCTTGACAAA 449
QY 21 ThMeSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 450 AGTATGAGTGAGATATCTCTCAGGTTTCAAGCTAGAAATGCCCTTCTTCAATCTT 509
QY 41 ProGlnAlaGlnLeuAlaSerIleTrrPamLeuSerAspIleAspGlnAspGlyLysLeu 60
Db 510 TCTCAACTGAGCTGCTACTATTGACTCTGCTGACGCTGATGCTGATGACAGCTA 569
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 570 AAAGCAGAAAGATTATTCTTGCATGACCTTACTGRCATGSCCAAGCTGACAGCCA 629
QY 81 LeuProValIleuProProGlnIlyrIleProProSerPheArg 95
Db 630 TTACCACTGACTTACCTCCTGAGCTTGTTCTTCCTCCATCTTTCAGG 674
```

```
RESULT 10
US-09-764-881-54
/ Sequence 54, Application US/09764881
/ Publication No. US20020086821A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PT207
/ CURRENT APPLICATION NUMBER: US/09/764,881
/ CURRENT FILING DATE: 2001-01-17
/ Prior application data removed - refer to PALM or file wrapper
/ NUMBER OF SEQ ID NOS: 192
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 54
/ LENGTH: 955
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (550)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (934)
```

```
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (942)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (944)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (948)
/ OTHER INFORMATION: n equals a,t,g, or c
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-764-881-54
```

```
Alignment Scores:
Pred. No.: 3.78e-38 Length: 955
Score: 334.00 Matches: 64
Percent Similarity: 81.05% Conservative: 13
Best Local Similarity: 67.37% Mismatches: 18
Query Match: 66.67% Indels: 0
DB: 13 Gaps: 0
```

US-09-720-934-2_COPY_215_310 (1-96) x US-09-764-881-54 (1-955)

```
QY 1 TrrAlvAlProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
Db 390 TGGGCGCTTCTAGCCTACAGATTAAATATCGCAAAATTTAACTCTTGACAAA 449
QY 21 ThMeSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 450 AGTATGAGTGAGATATCTCTCAGGTTTCAAGCTAGAAATGCCCTTCTTCAATCTT 509
QY 41 ProGlnAlaGlnLeuAlaSerIleTrrPamLeuSerAspIleAspGlnAspGlyLysLeu 60
Db 510 TCTCAACTGAGCTGCTACTATTGACTCTGCTGACGCTGATGCTGATGACAGCTA 569
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 570 AAAGCAGAAAGATTATTCTTGCATGACCTTACTGRCATGSCCAAGCTGACAGCCA 629
QY 81 LeuProValIleuProProGlnIlyrIleProProSerPheArg 95
Db 630 TTACCACTGACTTACCTCCTGAGCTTGTTCTTCCTCCATCTTTCAGG 674
```

```
RESULT 11
US-10-242-747-54
/ Sequence 54, Application US/10242747
/ Publication No. US20040005577A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PT207C1
/ CURRENT APPLICATION NUMBER: US/10/242,747
/ CURRENT FILING DATE: 2002-09-13
/ PRIOR APPLICATION NUMBER: 09/764,881
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: 60/179,065
/ PRIOR FILING DATE: 2000-01-31
/ PRIOR APPLICATION NUMBER: 60/180,628
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: 60/214,886
/ PRIOR FILING DATE: 2000-06-28
/ PRIOR APPLICATION NUMBER: 60/217,487
/ PRIOR FILING DATE: 2000-07-11
/ PRIOR APPLICATION NUMBER: 60/225,758
/ PRIOR FILING DATE: 2000-08-14
/ PRIOR APPLICATION NUMBER: 60/220,963
/ PRIOR FILING DATE: 2000-07-26
/ PRIOR APPLICATION NUMBER: 60/217,496
/ PRIOR FILING DATE: 2000-07-11
/ PRIOR APPLICATION NUMBER: 60/225,447
/ PRIOR FILING DATE: 2000-08-14
/ PRIOR APPLICATION NUMBER: 60/218,290
/ PRIOR FILING DATE: 2000-07-14
/ Remaining Prior Application data removed - See File Wrapper or PALM.
```


LENGTH: 3976
TYPE: DNA
ORGANISM: Aspergillus nidulans
US-10-369-493-36727

Alignment Scores:

Pred. No.:	1,37e-20	Length:	3976
Score:	218.50	Matches:	43
Percent Similarity:	62.89%	Conservative:	18
Best Local Similarity:	44.33%	Mismatches:	35
Query Match:	43.61%	Indels:	1
	16	Gaps:	1

US-09-720-934-2_COPY_215_310 (1-96) x US-10-369-493-36727 (1-3976)

QY 1 TYPAlaValProGlnSerSerArgLeuLysTYrArgGlnLeuPheAsnSerHisAspLys 20

Db 1043 TGGCTATTCTTCGCGAGAGAACTCCAGTTTGACAACTACTACTGTCGATACG 1102

QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40

Db 1103 ACGAATCTGGGTATTAACCGCGATCAAGCTGTACCGTTCTTACGAAGCGCGACGTA 1162

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPheLeuSerAspIleAspGlnAspGlyLysLeu 60

Db 1163 CCGAGAGAGACCTTGGCGCAGATTGGCATCTTGCGATTTGATGCTGATGGCCAGTTA 1222

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGln 79

Db 1223 ACAGGAGATGATGTTGGCGGTGGCATGTATTGATTCGCGAATTCACCAACAGAGGC 1282

QY 80 ProLeuProValLeuProProGluTyrIleProSerSerPheArgArg 96

Db 1283 CCCCTACCGCAGACTTACCAACCCGCTTGATACCCCAAGTATGGCGCA 1333

RESULT 14

US-09-884-441-60
Sequence 60, Application US/09884441
Patent No. US20020119158A1

GENERAL INFORMATION:

APPLICANT: Algate, Paul A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.462C7
CURRENT APPLICATION NUMBER: US/09/884,441
NUMBER OF SEQ ID NOS: 489
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 60
LENGTH: 480
TYPE: DNA
ORGANISM: Homo sapien
US-09-884-441-60

Alignment Scores:

Pred. No.:	2.41e-20	Length:	480
Score:	208.00	Matches:	44
Percent Similarity:	63.16%	Conservative:	16
Best Local Similarity:	46.32%	Mismatches:	33
Query Match:	41.52%	Indels:	2
	9	Gaps:	2

US-09-720-934-2_COPY_215_310 (1-96) x US-09-884-441-60 (1-480)

QY 1 TYPAlaValProGlnSerSerArgLeuLysTYrArgGlnLeuPheAsnSerHisAspLys 20

Db 56 TGGGCTATTACATCTGAAGACGTAAGTAAAGTAAACAGTTGATACCTC---AAA 112

QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40

Db 113 CCTCAGAGAGTTACATACAGGTATCAAGCCCGTACTTTTCTCAGTACAGGTCG 172

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPheLeuSerAspIleAspGlnAspGlyLysLeu 60

Db 173 CCGCCCCGGTATTAGCTGAATAATGGCCCTTATCATGATCTGAACAGATGGCAAGATG 232

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80

Db 233 GACCAAGCAAGATTTCTTATAGCTATGAATCAAGTTAAAGTTGCAAGGCCCAACAG 292

QY 81 LeuProProValLeuProPro---GluTyrIleProSerPhe 94

Db 293 CTGCCTGTAGTCTCCCTCATCATGAACAACCCCTATGTTTC 337

RESULT 15

US-09-907-969-60

Sequence 60, Application US/09907969
Publication No. US20030091580A1

GENERAL INFORMATION:

APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
APPLICANT: Algate, Paul A.
APPLICANT: Fling, Steven P.
APPLICANT: Retter, Marc W.
APPLICANT: Fanger, Gary Richard
APPLICANT: Reed, Steven G.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Hill, Paul
APPLICANT: Albane, Earl
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
AND DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.462C8
CURRENT APPLICATION NUMBER: US/09/907,969
NUMBER OF SEQ ID NOS: 596
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 60
LENGTH: 480
TYPE: DNA
ORGANISM: Homo sapiens
US-09-907-969-60

Alignment Scores:

Pred. No.:	2.41e-20	Length:	480
Score:	208.00	Matches:	44
Percent Similarity:	63.16%	Conservative:	16
Best Local Similarity:	46.32%	Mismatches:	33
Query Match:	41.52%	Indels:	2
	10	Gaps:	2

US-09-720-934-2_COPY_215_310 (1-96) x US-09-907-969-60 (1-480)

QY 1 TYPAlaValProGlnSerSerArgLeuLysTYrArgGlnLeuPheAsnSerHisAspLys 20

Db 56 TGGGCTATTACATCTGAAGACGTAAGTAAAGTAAACAGTTTATTAACCTC---AAA 112

QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40

Db 113 CCTCAGAGAGTTACATACAGGTATCAAGCCCGTACTTTTCTCAGTACAGGTCG 172

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrPheLeuSerAspIleAspGlnAspGlyLysLeu 60

Db 173 CCGGCCCCGGTATTAGCTGAATAATGGCCCTTATCATGATCTGAACAGATGGGAAGATG 232

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80

Db 233 GACCAAGCAAGATTTCTTATAGCTATGAATCAAGTTAAAGTTGCAAGGCCCAACAG 292

QY 81 LeuProProValLeuProPro---GluTyrIleProSerPhe 94

Db 293 CTGCCTGTAGTCTCCCTCATCATGAACAACCCCTATGTTTC 337

Search completed: July 1, 2004, 20:15:49

Fri Jul 2 10:23:38 2004

Job time : 300.52 secs

us-09-720-934-2_copy_215_310.rnpb

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:01:26 ; Search time 2887.87 Seconds
(without alignments)
992.694 Million cell updates/sec

Title: US-09-720-934-2_COPY_215_310
Perfect score: 501
Sequence: 1 WAVPOSSRLKXROLFNSHDK.....SGQPLPVLPPIPIPSFR 96

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q/cg2_1/USPTO.spool/p/US09720334/funat 30062004 064540 13442/app.query.fasta_1.1386
-DB=ESL -QFMT=fastap -SUFFIX=est -MIMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEASize=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09720334.@CGN_1_12431@funat 30062004 064540 13442 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-FAPEXT=7 -Ygapop=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_hum:*
19: em_gss_hum:*
20: em_gss_hum:*
21: em_gss_hum:*
22: em_gss_hum:*
23: em_gss_hum:*
24: em_gss_hum:*
25: em_gss_hum:*
26: em_gss_hum:*
27: em_gss_hum:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	501	100.0	625	14	CF532126	CF532126 UI-M-GHO-
2	501	100.0	641	12	BM728997	BM728997 UI-E-BOI-
3	501	100.0	729	13	BU704308	BU704308 UI-M-FOO-
4	501	100.0	729	14	CF742748	CF742748 UI-M-HBO-
5	501	100.0	732	14	CA750495	CA750495 UI-M-FYO-
6	501	100.0	979	13	CF538122	CF538122 UI-M-GIO-
7	501	100.0	979	13	BO719508	BO719508 AGENCOURT
8	501	100.0	2079	11	BC013578	BC013578 Homo sapi
9	501	100.0	2126	11	BC020269	BC020269 Homo sapi
10	501	100.0	5385	11	BC062938	BC062938 Mus muscu
11	492	98.2	490	10	AM370310	AM370310 RCL-BT025
12	485	96.8	842	9	AL870708	AL870708 AL870708
13	484	96.6	776	13	BU119880	BU119880 603142511
14	483	96.4	776	13	BK758765	BK758765 BX758765
15	483	96.4	853	13	BK695683	BK695683 BX695683
16	474	94.6	646	12	BG161771	BG161771 de42408.X
17	471	94.0	819	14	CF741757	CF741757 UI-M-HBO-
18	462	92.2	445	10	BB849620	BB849620 BB849620
19	462	92.2	548	10	AM581906	AM581906 MR4-ST011
20	460	91.8	765	14	CF539131	CF539131 UI-M-GIO-
21	458	91.4	359	13	BY220693	BY220693 BY220693
22	448	89.4	687	14	CB521525	CB521525 UI-M-GHO-
23	431	86.0	405	13	BY520359	BY520359 BY520359
24	425	84.8	907	13	BU128229	BU128229 603114019
25	421	84.0	585	14	CB518599	CB518599 UI-M-GHO-
26	418	83.4	486	12	BM152392	BM152392 TCBPAP188
27	417.5	83.3	1201	9	AL531589	AL531589 AL531589
28	417	83.2	606	10	AM910367	AM910367 UT80908.Y
29	405	82.8	722	14	CF533008	CF533008 UI-M-FYO-
30	407	81.2	911	13	BQ942708	BQ942708 AGENCOURT
31	376	75.0	724	10	BE373101	BE373101 601224527
32	352	70.3	886	13	BU17489	BU17489 603849519
33	349	69.7	425	9	AA154049	AA154049 m668405.X
34	349	69.7	639	10	BB242625	BB242625 BB242625
35	349	69.7	861	14	CD653097	CD653097 AGENCOURT
36	349	69.7	1550	11	AK042449	AK042449 Mus muscu
37	349	69.7	2352	11	AK034367	AK034367 Mus muscu
38	349	69.7	2641	11	AK029325	AK029325 Mus muscu
39	348	69.5	897	13	BO735187	BO735187 AGENCOURT
40	342	68.3	783	12	BG107389	BG107389 602280786
41	337	67.3	546	10	AW668653	AW668653 111193 MA
42	331	66.1	395	10	AW889826	AW889826 R55-NT003
43	331	66.1	574	10	BF816839	BF816839 MR2-CI012
44	316	63.1	627	9	AV896004	AV896004 AV896004
45	316	63.1	636	13	BM121860	BM121860 BM121860

ALIGNMENTS

RESULT 1
LOCUS CF532126 625 bp mRNA linear EST 12-SEP-2003
DEFINITION UI-M-GHO-cgu-d-05-0-UI.r1 NIH_BMAP_GHO Mus musculus cDNA clone
IMAGE:30356524 5', mRNA sequence.
ACCESSION CF532126
VERSION CF532126.1 GI:34584094
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 625)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouse1.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: PYX-5.

FEATURES
 source

Location/Qualifiers
 1. 625
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:30356524"
 /tissue_type="whole brain"
 /dev_stage="1, 5, and 15 days newborn"
 /lab_host="DH10B (TI phage resistant)"
 /clone_id="NIH_BMAP_GH0"
 /note="Organ: Brain; Vector: PYX-Asc; Site: 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into PYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAAGTGAAT. This library was created for the University Iowa Brain Anatomy Project (BMAP). Gene Discovery in the Developing Mouse Nervous System", supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
 Pred. No.: 8.35e-47 Length: 625
 Score: 501.00 Matches: 96
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x CFS32126 (1-625)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAnsSerHisAspLys 20
 Db 42 TGGGCTGTGCTCGATCATCAAGGCTGAATACAGGCGATTATCAACAGCCACAGCAAA 101
 QY 21 ThMetSerGlyHisLysThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIleu 40
 Db 102 ACTATGATGTGACACTTACAGCTCCCGAGCAAGAACTATTCATGCAATCAAGTTTA 161
 QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
 Db 162 CCCGAGGCTCGAGTGGCTTCAATATGAAATCTTCTGACATTGATCAAGATGAAAACTC 221
 QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
 Db 222 ACTGCAAGAAGATTATCTAGCTTAGACCCMAATTGATGTGCAATGCTGTCGACAGCA 281
 QY 81 LeuProValLeuLeuProProGlnTyrIleProProSerPheArgArg 96
 Db 282 CTGGCGGCGCTCTCTCCCTCCAGAAATACATCCCTCTTCTTCAGAAAG 329

RESULT 2

BM728997
 LOCUS 641 bp mRNA linear EST 01-MAR-2002
 DEFINITION UI-E-B01-aiw-e-23-0-UI.r1 UI-E-B01 Homo sapiens cDNA clone
 accession BM728997
 VERSION UI-E-B01-aiw-e-23-0-UI 5', mRNA sequence.
 KEYWORDS
 SOURCE EST.
 ORGANISM Homo sapiens (human)

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 TITLE 1 (bases 1 to 641)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

FEATURES
 source

Location/Qualifiers
 1. 641
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-B01-aiw-e-23-0-UI"
 /tissue_type="fetal eye"
 /dev_stage="fetal"
 /lab_host="DH10B (Life Technologies) (TI phage resistant)"
 /clone_id="UI-E-B01"
 /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-B01 is a normalized cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGCGTATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

Alignment Scores:
 Pred. No.: 8.65e-47 Length: 641
 Score: 501.00 Matches: 96
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x BM728997 (1-641)

QY 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAnsSerHisAspLys 20
 Db 75 TGGGCTGTGCTCGATCATCAAGCTGAATACAGCAATTATTCATGATGACAGCAAA 134

QY 21 ThMeSerGIyhiSteuThrglyProGlnAlaArgThrIleLeuMetGlnSerIeu 40
 |||
 DB 135 ACTATAGTGGACACTTAACAGGTCCTCCCAAGCAAGAACTATCTTATGCAAGTCAAGTTTA 194
 QY 41 ProGlnAlaGlnLeuAlaSerIleTPanLeuSerAspIleAspGlnAspGlyIysIeu 60
 |||
 DB 195 CCACAGGCTCAGCTGGCTTCAATATGGAAATCTTTGACATTCATCAAGATGGAAAACTT 254
 QY 61 ThrAlaGlnGluPheIleuAlaMetHisLeuIleAspValAlaMetSerGIyGlnPro 80
 |||
 DB 255 ACAGCAGAGAAATTATCTCGGCAATGCACTCATGATGATAGTCTGTGCGCAACCA 314
 QY 81 LeuProProValLeuProProGIyTyrIleProProSerPheArgArg 96
 |||
 DB 315 CTGCCACCTGTCTGCTGCTCCAGAAATCATTCACCTTTTGAAGA 362

RESULT 3
 BUT04308 729 bp mRNA linear EST 15-JUL-2003
 LOCUS UI-M-F00-bzs-j-21-0-UI.r1 NIH_BMAP_F00 Mus musculus cDNA clone
 DEFINITION IMAGE:6406772 5', mRNA sequence.
 ACCESSION BUT04308
 VERSION BUT04308.1 GI:23632308
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgc.nci.nih.gov/
 1 (bases 1 to 729)
 REFERENCE National Institutes of Health, Mammalian Gene Collection (MGC)
 AUTHORS Unpublished (1999)
 TITLE Contact: Robert Strausberg, Ph.D.
 JOURNAL Email: cgabbs-remail.nih.gov
 COMMENT Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seg primer: PYX-5.
 Location/Qualifiers
 1. 729
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6406772"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_id="NIH_BMAP_F00"
 /note="Organ: Brain; Vector: PYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into PYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is TGAAGAGGCC. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 1.04e-46 Length: 729
 Score: 501.00 Matches: 96
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) x BUT04308 (1-729)

QY 1 TriPalAvalProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
 |||
 DB 259 TGGGCTGTGCTCCTCAGTCATCAAGCTGAATACAGGACAGTATTCACAGCCACGACAA 318
 QY 21 ThMeSerGIyhiSteuThrglyProGlnAlaArgThrIleLeuMetGlnSerIeu 40
 |||
 DB 319 ACTATAGTGGACACTTAACAGGTCCTCCCAAGCAAGAACTATCTTCAATGCAATCAAGTTTA 378
 QY 41 ProGlnAlaGlnLeuAlaSerIleTPanLeuSerAspIleAspGlnAspGlyIysIeu 60
 |||
 DB 379 CCCACAGCTCAGCTGGCTTCAATATGGAAATCTTTGACATTCATCAAGATGGAAAACTC 438
 QY 61 ThrAlaGlnGluPheIleuAlaMetHisLeuIleAspValAlaMetSerGIyGlnPro 80
 |||
 DB 439 ACTGACAGAGAAATTATCTCGGCAATGCACTCATGATGATGCTGCTGCGCAACCA 498
 QY 81 LeuProProValLeuProProGIyTyrIleProProSerPheArgArg 96
 |||
 DB 499 CTGCCACCTGTCTGCTGCTCCAGAAATCATTCACCTTTTGAAGA 546

RESULT 4
 CF742748 729 bp mRNA linear EST 10-OCT-2003
 LOCUS UI-M-HB0-clm-c-18-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone
 DEFINITION IMAGE:30617825 5', mRNA sequence.
 ACCESSION CF742748
 VERSION CF742748.1 GI:37639087
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgc.nci.nih.gov/
 1 (bases 1 to 729)
 REFERENCE National Institutes of Health, Mammalian Gene Collection (MGC)
 AUTHORS Unpublished (1999)
 TITLE Contact: Robert Strausberg, Ph.D.
 JOURNAL Email: cgabbs-remail.nih.gov
 COMMENT Tissue Procurement: Dr. James Lin University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/mousefl.html
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 The following repetitive elements were found in this cDNA
 sequence: 589-705, >(CAG)n#Simple_repeat
 Seg primer: PYX-5.
 Location/Qualifiers
 1. 729
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:30617825"
 /tissue_type="whole eye"
 /dev_stage="embryo 12.5,13.5,14.5 dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_id="NIH_BMAP_HB0"
 /note="Organ: Eye; Vector: PYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

FEATURES
 source

1. 729
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:30617825"
 /tissue_type="whole eye"
 /dev_stage="embryo 12.5,13.5,14.5 dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_id="NIH_BMAP_HB0"
 /note="Organ: Eye; Vector: PYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TTTATGAGT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

ORIGIN

Alignment Scores:

Pred. No.:	1,04e-46	Length:	729
Score:	501.00	Matches:	96
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

US-09-720-934-2_COPY_215_310 (1-96) x CF742748 (1-729)

QY 1 TTPALAVAlProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
 Db 116 TGGGCTGGCTCAGTCATCAAGGCTGAATATACAGGAGATTATTCACAGCAGACACAA 175

QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
 Db 176 ACTATGATGAGACACTTAACAGGTCCCGAGCAAGAACTATTCATCAATCAAGTTTA 235

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysIeu 60
 Db 236 CCCAGGCTCAGCTGGCTTCATATGAGATCTTTCTGACATTGATCAAGATGAGAAATCTC 295

QY 61 ThrAlaGlnIupheIleuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
 Db 296 ACTGAGAGAGATTATTCCTAGTATGACCACTAATGATTGTCATGTCGTGTCAGCCA 355

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArgArg 96
 Db 356 CTGGCCGCCGCTGCTCCCTCCAGAAATACATCCCTCTCTCTTCAGAGAA 403

ORIGIN

Alignment Scores:

Pred. No.:	1.05e-46	Length:	732
Score:	501.00	Matches:	96
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

US-09-720-934-2_COPY_215_310 (1-96) x CA750495 (1-732)

QY 1 TTPALAVAlProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
 Db 114 TGGGCTGGCTCAGTCATCAAGGCTGAATATACAGGAGATTATTCACAGCAGACACAA 173

QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
 Db 174 ACTATGATGAGACACTTAACAGGTCCCGAGCAAGAACTATTCATCAATCAAGTTTA 233

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysIeu 60
 Db 234 CCCAGGCTCAGCTGGCTTCATATGAGATCTTTCTGACATTGATCAAGATGAGAAACTC 293

QY 61 ThrAlaGlnIupheIleuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
 Db 294 ACTGAGAGAGATTATTCCTAGTATGACCACTAATGATTGTCATGTCGTGTCAGCCA 353

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArgArg 96
 Db 354 CTGGCCGCCGCTGCTCCCTCCAGAAATACATCCCTCTCTCTTCAGAGAA 401

RESULT 5
 CA750495 732 bp mRNA linear EST 09-JUL-2003
 LOCUS UI-M-FY0-cde-p-10-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
 DEFINITION IMAGE:6832091 5', mRNA sequence.
 ACCESSION CA750495
 VERSION CA750495.1 GI:25575732
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 732)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/IML at: http://image.llnl.gov

FEATURES

source

Location/Qualifiers

1..732

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:6832091"

/tissue_type="whole brain"

/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_id="NIH_BMAP_FY0"

/note="Organ: Brain; Vector: pX-Asc, Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.:	1.05e-46	Length:	732
Score:	501.00	Matches:	96
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

US-09-720-934-2_COPY_215_310 (1-96) x CA750495 (1-732)

QY 1 TTPALAVAlProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
 Db 114 TGGGCTGGCTCAGTCATCAAGGCTGAATATACAGGAGATTATTCACAGCAGACACAA 173

QY 21 ThrMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
 Db 174 ACTATGATGAGACACTTAACAGGTCCCGAGCAAGAACTATTCATCAATCAAGTTTA 233

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysIeu 60
 Db 234 CCCAGGCTCAGCTGGCTTCATATGAGATCTTTCTGACATTGATCAAGATGAGAAACTC 293

QY 61 ThrAlaGlnIupheIleuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
 Db 294 ACTGAGAGAGATTATTCCTAGTATGACCACTAATGATTGTCATGTCGTGTCAGCCA 353

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArgArg 96
 Db 354 CTGGCCGCCGCTGCTCCCTCCAGAAATACATCCCTCTCTCTTCAGAGAA 401

RESULT 6
 CF538122 736 bp mRNA linear EST 12-SEP-2003
 LOCUS UI-M-G10-cho-b-09-0-UI.r1 NIH_BMAP_G10 Mus musculus cDNA clone
 DEFINITION IMAGE:30536000 5', mRNA sequence.
 ACCESSION CF538122
 VERSION CF538122.1 GI:34590104
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 736)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)

The following repetitive elements were found in this cDNA sequence: 637-722, >(CAG)n#simple_repeat
 Seq primer: pX-5.

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: PYX-5.
FEATURES
Location/Qualifiers
1..736
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30536000"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_1lb="NIH BMAP G10"
/note="Organ: Brain; Vector: PYX-Asc; Site 1: Bcor I;
Site 2: Not I; The library was constructed according
to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction. Ligated
with Bcor I adaptor, digested with NotI and then cloned
directionally into PYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemlin Chin, Ph.D.,
program coordinator."

ORIGIN
Alignment Scores:
Pred. No.: 1,05e-46 Length: 736
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-09-720-934-2_COPY_215_310 (1-96) x CFS38122 (1-736)
QY 1 TTPAlaValProGlnSerSerArgLeuLYTYrArgGlnLeuPheAnsSerHisAspLYs 20
Db 185 TGGGCTGTGCTCAGTCATCAAGCGTGAATATACAGGACGATTATCAACCCGACCA 244
QY 21 ThMeSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 245 ACTATAGTGGACACTTAACAGGTCGCCAGGCAAGAACTATTCTCATGCAATCAAGTTTA 304
QY 41 ProGlnAlaGlnLeuAlaSerIleTTPAsnLeuSerAspIleAspGlnAspGlyLYsLeu 60
Db 305 CCCCAAGCTCAGCTGGCTTCAATATGGAATCTTTGCACATGATCAAGATGAAAACTC 364
QY 61 ThrAlaGlnGluIleuPheLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 365 ACTGCAGAGAAATTTATCTCAGCTATGCACTCAATGATGTCCTGCTGGTCAGCCA 424
QY 81 LeuProValLeuPheProGluTYrIleProProSerPheArgArg 96
Db 425 CTGCGCGCGGTCTGCTCCAGAAATACATCCCTCTCTTCAGAGAA 472
RESULT 7
BQ719508
LOCUS BQ719508 979 bp mRNA linear EST 16-JUL-2002

DEFINITION AGENCOURT 8219793 lupski sympathetic_trunk Homo sapiens cDNA clone
IMAGE:6188111 5', mRNA sequence.
ACCESSION BQ719508
VERSION BQ719508.1 GI:21858405
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
1 (bases 1 to 979)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLM13583 row: c column: 24
High quality sequence stop: 643.
FEATURES
Location/Qualifiers
1..979
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6188111"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/note="Vector: PCWV-SPORE (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGTCG-3' and
5'-GACTAGTCTAGATCGGAGCGGCGCCCTT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

ORIGIN
Alignment Scores:
Pred. No.: 1,58e-46 Length: 979
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-09-720-934-2_COPY_215_310 (1-96) x BQ719508 (1-979)
QY 1 TTPAlaValProGlnSerSerArgLeuLYTYrArgGlnLeuPheAnsSerHisAspLYs 20
Db 353 TGGGCTGTGCTCAGTCATCAAGCTGAATAACAGCAATTAATCAATGATGACAA 412
QY 21 ThMeSerGlyHisLeuThrglyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 413 ACTATAGTGGACACTTAACAGGTCCCAAGCAAGAACTATTCTTATGCAGTCAAGTTTA 472
QY 41 ProGlnAlaGlnLeuAlaSerIleTTPAsnLeuSerAspIleAspGlnAspGlyLYsLeu 60
Db 473 CCACAGGCTCAGCTGGCTTCAATATGGAATCTTTGCACATGATCAAGATGAAAACTT 532
QY 61 ThrAlaGlnGluIleuPheLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 533 ACAGCAGAGAAATTTATCTCAGCAATGCACTCATGATGATGATGTCGCGCAACA 592

Qy 81 LeuProProValLeuProProGluTyrIleProProSerPheArgArg 96
Db 593 CTGCACCTGCTCTGCTCCAGAAATACATTCACCTTTTAAAGA 640

RESULT 8
LOCUS BC013578
DEFINITION Homo sapiens, similar to intersectin 1 (SH3 domain protein), clone IMAGE:3878242, mRNA.
ACCESSION BC013578
VERSION BC013578.1 GI:15488896
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2079)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (04-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.sbgc.stanford.edu>
Contact: (Dickson, Mark) mcddapaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

FEATURES
source 1..2079
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3878242"
/tissue_type="lung, large cell carcinoma"
/clone_id="NIH MGC_68"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN
Alignment Scores:
Pred. No.: 4.64e-46 Length: 2079
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) * BC013578 (1-2079)

Qy 1 TTPAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
Db 774 TGGGCTGTCCTCAGTCATCAAGACTGAATACAGGCGCAATATTCATATGATCAGACAAA 833

Qy 21 ThMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 834 ACTATGATGGACACTTAACAGATTCCTCCACAGCAAGAACTATTTATGCAAGTCAAGTTTA 893

Qy 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysLeu 60

Db 894 CCACAGGCTCAGCTGGCTTCATATGCAATCTTTCTACATTTGATCAAGATGAAAACTT 953

Qy 61 ThrAlaGluGlnPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 954 ACACGACAGGAATTTATCTTGCAATGACACTGATGATGAGCTATGCTGGCAACCA 1013

Qy 81 LeuProProValLeuProProGluTyrIleProProSerPheArgArg 96
Db 1014 CTGCACCTGCTCTGCTCCAGAAATACATTCACCTTTTAAAGA 1061

RESULT 9
LOCUS BC020269
DEFINITION Homo sapiens, clone IMAGE:4899011, mRNA.
ACCESSION BC020269
VERSION BC020269.1 GI:17939664
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2126)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca

FEATURES
source 1..2126
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4899011"
/tissue_type="Pancreas, epithelioid carcinoma"
/clone_id="NIH MGC_42"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

ORIGIN
Alignment Scores:
Pred. No.: 4.79e-46 Length: 2126
Score: 501.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRAL Plate: 40 Row: n Column: 1
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4504796
This clone has the following problem: no cloning site / microdeletion.

REMARK	COMMENT
NIH-MGC Project URL: http://mgc.ncl.nih.gov	
Contact: MGC help desk	
Email: cgabs-remail.nih.gov	
Tissue Procurement: Dr. James Lin, University of Iowa	
CNA Library Preparation: M. Bento Soares, University of Iowa	
CNA Library Arrayed by: The I.M.A.G.E. Consortium (IHLI)	
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.	

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (pages 1 to 490)
HCGP <http://www.ludwig.org.br/ORESTES>.
The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao paulo-SP,

Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l1=RC1&t2=RC1-BT0255-041059-011-c09&t3=1999-10-04&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 24
 High quality sequence stop: 489.
 Location/Qualifiers
 1. 490
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_1lb="BT0255"
 /note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
 Alignment Scores:
 Pred. No.: 6,26e-46 Length: 490
 Score: 492.00 Matches: 95
 Percent Similarity: 98.96% Conservative: 0
 Best Local Similarity: 98.96% Mismatches: 1
 Query Match: 98.20% Indels: 0
 DB: 10 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) X AM370310 (1-490)

QY 1 TrrAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
 Db 337 TGGGCTGTCTCTCGATCAAGACCTGAAATACAGGCAATTATTCATAGTCATGACAA 278

QY 21 ThMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
 Db 277 ACTATGATGGACACTTAACAGGTCCTCCCAAGCAAGAACTATTTATGCAAGTCAAGTTTA 218

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysIeu 60
 Db 217 CCACAGGCTCAGCTGGCTCAATATGGAATCTTTCTGACATTTGTTCAAGATGGAATAACTT 158

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
 Db 157 ACMGCAAGAAATTTATTCCTGGCAATGACCTCATTTGATGCTATGCTCTGGCCAAACC 98

QY 81 LeuProBovValLeuProProGluTyrIleProBosSerPheArgArg 96
 Db 97 CTGCCACCTGTCTGCTGCCAGAAATACCTTCACCTTTTAAAGA 50

RESULT 12
 AL870708 842 bp mRNA linear EST 03-DEC-2003
 LOCUS AL870708 XGC-egg Silurana tropicalis cDNA clone TEG9141p04 5', mRNA
 DEFINITION sequence.
 ACCESSION AL870708 GI:38665439
 VERSION AL870708.2 GI:38665439
 KEYWORDS EST.
 SOURCE Silurana tropicalis (western clawed frog)
 ORGANISM Silurana tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae, Pipridae; Xenopodinae; Silurana.
 1 (bases 1 to 842)
 Crounig,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
 Sanger Xenopus tropicalis EST project 2001 (11_2003)

REFERENCE
 AUTHORS
 TITLE

JOURNAL
 COMMENT Unpublished (2003)
 On Sep 15, 2002 this sequence version replaced gi:22890973.
 Contact: Taylor R
 Sanger Institute
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 Sanger Xenopus tropicalis EST project 2001
 TROPICALIS_SEQUENCE_ID: TEG9141p04.plkBP6
 Sequencing primer: SP6
 This sequence is from a Xenopus Gene Collection (XGC) library
 constructed by Aaron M. Zorn.
 cDNA was oligo dt primed from 5ug of poly A+ RNA from egg.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end.
 Vector: pCS107; Site 1: EcoRI; Site 2: NotI
 Host: Escherichia coli XL1-blue.
 Location/Qualifiers
 1. 842
 /organism="Silurana tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="TEG9141p04"
 /dev_stage="egg"
 /lab_host="Escherichia coli XL1-blue"
 /clone_1lb="XGC-egg"
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dt primed from 5ug of poly A+ RNA from egg. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end"

ORIGIN
 Alignment Scores:
 Pred. No.: 8.5e-45 Length: 842
 Score: 485.00 Matches: 92
 Percent Similarity: 98.96% Conservative: 3
 Best Local Similarity: 95.83% Mismatches: 1
 Query Match: 96.81% Indels: 0
 DB: 9 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) X AL870708 (1-842)

QY 1 TrrAlaValProGlnSerSerArgLeuLysTyrArgGlnLeuPheAsnSerHisAspLys 20
 Db 446 TGGGCTGTACACAGCTTCAAGGTGAAGTACAGACAGTTGTTCAACAGCCATGACAA 505

QY 21 ThMetSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerIeu 40
 Db 506 ACCATGATGGAGCAATTTAAACAGGCTCAAGCAAGAACTATTTATGCAAGTCAAGTTTA 565

QY 41 ProGlnAlaGlnLeuAlaSerIleTyrAsnLeuSerAspIleAspGlnAspGlyLysIeu 60
 Db 566 CCGAGAGCCAGCTAGCTAGCATGATGGAACCTTTCAACATTTGACCTGATGAGGAAAACTA 625

QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
 Db 626 ACGGCCAAGAAATTTATTTAGCTATGACACTTAATATGATGGCCATGCTGGCCAGCCA 685

QY 81 LeuProBovValLeuProProGluTyrIleProBosSerPheArgArg 96
 Db 686 CTCTCTCTGTTCTGCTCTCGAGTATATTCCTCCCATCTTTAAAGA 733

RESULT 13
 BU119880 786 bp mRNA linear EST 25-NOV-2002
 LOCUS BU119880
 DEFINITION 603142511F1 CSEBCH16 Gallus gallus cDNA clone CHES1136f4 5', mRNA
 sequence.
 ACCESSION BU119880
 VERSION BU119880.1 GI:25329000
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

RESULT 15
BX695683/c

LOCUS BX695683 XGC-neurula Silurana tropicalis cDNA clone TNeu1ln07 3',
DEFINITION mRNA sequence.
ACCESSION BX695683
VERSION BX695683.1 GI:38357890
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 853)
REFERENCE Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
AUTHORS Sanger Xenopus tropicalis EST project 2001 (11_2003)
TITLE Unpublished (2003)
JOURNAL Contact: Croning MDR
COMMENT Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: tropesanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TNeu1ln07.q1xT7
Sequencing primer: T7
This sequence is from a
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site_1: EcoRI; Site_2: NotI
Host: Escherichia coli DH10B.
Location/Qualifiers
1..853
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TNeu1ln07"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/clone_lib="XGC-neurula"
/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

Alignment Scores:
Pred. No.: 1.46e-44 Length: 853
Score: 483.00 Matches: 92
Percent Similarity: 98.95% Conservative: 2
Best Local Similarity: 96.84% Mismatches: 1
Query Match: 96.41% Indels: 0
DB: 13 Gaps: 0

US-09-720-934-2_COPY_215_310 (1-96) * BX695683 (1-853)

QY 1 TnpAlaValProGlnSerSerAkgLeuYstYrArgGlnLeuPheAsnSerHisAspLys 20
Db 547 TGGGCTGTACCAAGTCTTCAAGTTGAGTACAGACAGTTGTCACAGCCACGACAA 488
QY 21 ThmWtSerGlyHisLeuThrGlyProGlnAlaArgThrIleLeuMetGlnSerSerLeu 40
Db 487 ACCATGAGTGGGCACTTAACAGGTCCTCAAGCAGAACTATTCTTATGCACTCAAGTTTA 428
QY 41 ProGlnAlaGlnLeuAlaSerIleTrpAsnLeuSerAspIleAspGlnAspGlyLysLeu 60
Db 427 CCGAGAGCCAGCTAGCTACCATATGGAACCTTCAGACATTGACCTAGATGAGAAAACTA 368
QY 61 ThrAlaGlnGluPheIleLeuAlaMetHisLeuIleAspValAlaMetSerGlyGlnPro 80
Db 367 ACGGCCAAGATTTATTAGTTATGACATTATATAGATGTGGCAATGTCTGGCAGCCA 308

QY 81 LeuProProValLeuProProGluTyrIleProProSerPheArg 95
Db 307 CTTCTCCAGTTCTGCTCTCTGAGTATATTCCTCCATCTTTTAGA 263

Search completed: July 1, 2004, 19:48:34
Job time : 2895.87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:03:00 ; Search time 42.7143 Seconds
(without alignments)
792.522 Million cell updates/sec

Title: US-09-720-934-2_COPY_740_800
Perfect score: 333
Sequence: 1 VAVVYRRALYPERSHDEI.....GGELKKGKGFPPANYAEKIP 61

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool/p/US09720934/funat.30062004.064540.13455/app.query.fasta_1.1386
-DB=Issued Patents NA -QFMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPOPT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09720934.@CEN 1.1.284.@runat.30062004.064540.13455 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUTERY -NEG SCORES=0 -WAIT -DSBLOC=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: Issued Patents NA.*
2: /cgn2_6/prodata/2/ina/5A COMB.seq.*
3: /cgn2_6/prodata/2/ina/5B COMB.seq.*
4: /cgn2_6/prodata/2/ina/6A COMB.seq.*
5: /cgn2_6/prodata/2/ina/6B COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	215	64.6	2873	4	US-08-630-915A-193
2	140	42.0	970	4	US-09-833-381-1294
3	129	38.7	1392	1	US-08-475-894-3
4	129	38.7	1392	1	US-08-484-710-3
5	129	38.7	1392	2	US-08-484-709-3
6	129	38.7	1392	2	US-08-474-697-3
7	129	38.7	1659	1	US-08-475-894-1
8	129	38.7	1659	1	US-08-484-710-1
9	129	38.7	1659	2	US-08-484-709-1
10	129	38.7	1659	3	US-08-474-697-1
11	129	38.7	3143	4	US-08-671-354-1
12	128.5	38.6	2757	1	US-08-306-691B-48

13	128.5	38.6	2757	5	PCT-US93-06251-79	Sequence 79, Appl
14	124.5	37.4	2793	1	US-07-646-537B-1	Sequence 1, Appl
15	117	35.1	1277	4	US-08-630-915A-7	Sequence 7, Appl
16	116	34.8	691	4	US-08-630-915A-220	Sequence 220, Appl
17	116	34.8	901	5	PCT-US96-08950-1	Sequence 1, Appl
18	116	34.8	901	5	PCT-US96-09127-1	Sequence 1, Appl
19	116	34.8	1133	4	US-08-630-915A-35	Sequence 35, Appl
20	115.5	34.7	3111	4	US-10-014-882-1	Sequence 1, Appl
21	115.5	34.7	3518	4	US-10-014-882-3	Sequence 3, Appl
22	115	34.5	5325	4	US-10-164-595-17	Sequence 17, Appl
23	115	34.5	5334	4	US-10-164-595-21	Sequence 21, Appl
24	115	34.5	5581	4	US-10-164-595-19	Sequence 19, Appl
25	110	33.0	1661	2	US-08-815-176-2	Sequence 2, Appl
26	110	33.0	1661	4	US-09-197-344-2	Sequence 2, Appl
27	110	33.0	2100	3	US-08-938-830-2	Sequence 2, Appl
28	110	33.0	2100	3	US-09-020-222-2	Sequence 2, Appl
29	108	32.4	482	3	US-09-621-976-3867	Sequence 3867, Ap
30	108	32.4	3846	2	US-08-691-814B-3	Sequence 3, Appl
31	106	31.8	1199	4	US-08-630-915A-15	Sequence 15, Appl
32	106	31.8	1349	4	US-09-023-655-1176	Sequence 1176, Ap
33	106	31.8	1428	4	US-09-710-693-18	Sequence 18, Appl
34	106	31.8	1458	2	US-08-942-423-68	Sequence 68, Appl
35	106	31.8	1613	3	US-08-938-830-28	Sequence 28, Appl
36	106	31.8	1803	4	US-09-006-428A-18	Sequence 18, Appl
37	106	31.8	1858	4	US-09-006-428A-16	Sequence 16, Appl
38	106	31.8	2003	2	US-08-942-423-1	Sequence 1, Appl
39	106	31.8	2003	4	US-08-630-915A-25	Sequence 25, Appl
40	105	31.5	1382	4	US-09-820-005-1	Sequence 1, Appl
41	104.5	31.4	1240	4	US-09-833-381-888	Sequence 888, App
42	104	31.2	1636	4	US-08-630-915A-31	Sequence 31, Appl
43	104	31.2	1667	4	US-08-630-915A-13	Sequence 13, Appl
44	103.5	31.1	4762	4	US-09-300-958A-30	Sequence 30, Appl
45	102	30.6	4091	4	US-08-630-915A-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-08-630-915A-193
Sequence 193, Application US/08630915A
Patent No. 6309820
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOMLAKES, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,915A
FILING DATE: 03-Apr-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 193:
SEQUENCE CHARACTERISTICS:
LENGTH: 2873 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-08-630-915A-193

Alignment Scores:

Pred. No.: 8.52e-21
Score: 215.00
Percent Similarity: 74.14%
Best Local Similarity: 62.07%
Query Match: 64.56%
DB: 4

Length: 2873
Matches: 36
Conservative: 7
Mismatches: 15
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-08-630-915A-193 (1-2873)

QY 4 ValYTYTARGAlaLeuTYrProPheGluSerHisaspGluileThrIleGln 23

Db 231 GTGAATTATAGAGCATTAATACCTTTGAACAGCAAGACCATGATGAGATGATTAAAT 290

QY 24 ProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTYrIleuGly 43

Db 291 TCTGGAGATATTAATTCACGTTGATGAAAAACCGTAGAGAACTGGTTGCTTTATGCT 350

QY 44 GluLeuLysGlyLysThrGlyTYrPheProAlaAsnTYrAlaGluLysIlePro 61

Db 351 AGTTTCAAGAAATTTTGCTGCTTCATGCAATTATGTAGAAAAAATGCCA 404

RESULT 2

US-09-833-381-1294/C
Sequence 1294, Application US/0983381

Patent No. 6672186

GENERAL INFORMATION:

APPLICANT: Robison, Keith E.

TITLE OF INVENTION: NO. 6672186e1 Nucleic Acid and Protein Homologs

FILE REFERENCE: 5800-119

CURRENT APPLICATION NUMBER: US/09/833,381

PRIOR FILING DATE: 2001-04-11

PRIOR FILING DATE: 2000-02-29

NUMBER OF SEQ ID NOS: 2050

SOFTWARE: PASTESEQ for Windows Version 3.0

SEQ ID NO 1294

LENGTH: 970

TYPE: DNA

ORGANISM: Homo sapiens

US-09-833-381-1294

Alignment Scores:

Pred. No.: 1.01e-10

Score: 140.00

Percent Similarity: 67.27%

Best Local Similarity: 47.27%

Query Match: 42.04%

Length: 970
Matches: 26
Conservative: 11
Mismatches: 16
Indels: 2
Gaps: 1

US-09-720-934-2_COPY_740_800 (1-61) x US-09-833-381-1294 (1-970)

QY 7 ArgAlaLeuTYrProPheGluSerHisaspGluileThrIleGlnProGlyAsp 26

Db 376 AAGGCTTGTATTTTTCAGCGAAGATATAGTTGAATTTGAAAAAGGTGAC 317

QY 27 IleValMetValAspGluSerGlnThrGlyGluProGlyTYrIleuGlyGluLys 46

Db 316 ATTGTGATTATACGAGAAAAA-----GAAGAAGATGGTGGTTTGATCTTTGAAT 263

QY 47 GlyLysThrGlyTYrPheProAlaAsnTYrAlaGluLysIlePro 61

Db 47 GlyLysThrGlyTYrPheProAlaAsnTYrAlaGluLysIlePro 61

Db 262 GGGAAAAAGCCATTTTCTGCGCTTATGTGAGAGATTACT 218

RESULT 3

US-08-475-894-3
Sequence 3, Application US/08475894

Patent No. 5641748

GENERAL INFORMATION:

APPLICANT: Yen-Ming Hsu

TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESS: LAMIVE & COCKFIELD

STREET: 60 State Street, Suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/475,894

FILING DATE:

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Louis Myers

REGISTRATION NUMBER: 35,965

REFERENCE/DOCKET NUMBER: BGP-191

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1392 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-475-894-3

Alignment Scores:

Pred. No.: 5.98e-09

Score: 129.00

Percent Similarity: 64.91%

Best Local Similarity: 38.60%

Query Match: 38.74%

Length: 1392
Matches: 22
Conservative: 15
Mismatches: 20
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-08-475-894-3 (1-1392)

QY 5 TyTYrArgAlaLeuTYrProPheGluSerHisaspGluileThrIleGlnPro 24

Db 760 TACTGCAAGATTAATTTCCATATGAGCAGACAGATGATGATGACAAATCAAGAA 819

QY 25 GlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTYrIleuGly 44

Db 820 GGAGATATAGTCACTGCATCAATTAAGACTGCATCGACGATGAGGAGAGAG 879

QY 45 LeuLysGlyLysThrGlyTYrPheProAlaAsnTYrAlaGluLysIlePro 61

Db 880 CTGAACGCGACAGAGCGGTGTTCCCGATTAAGTCTGTAAGTACTTCCA 930

RESULT 4

US-08-484-710-3

Sequence 3, Application US/08484710

Patent No. 5656438

GENERAL INFORMATION:

APPLICANT: Yen-Ming Hsu

TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:


```
ADDRESS: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,710
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1392 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-484-710-3

Alignment Scores:
Pred. No.: 5.98e-09      Length: 1392
Score: 129.00           Matches: 22
Percent Similarity: 64.91% Conservative: 15
Best Local Similarity: 38.60% Mismatches: 20
Query Match: 38.74%      Indels: 0
DB: 1                   Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-08-484-710-3 (1-1392)
QY 5 TyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThrIleGlnPro 24
   ||| ::::::::::::::::::::||| ::::::::::::::::::::||| ::::::::::::::::::::|||
Db 760 TACTGCAAGATATATTTCATATGAGGACAGAAATGATGATTAATTGACAAATCAAGAA 819
   ||| ::::::::::::::::::::||| ::::::::::::::::::::||| ::::::::::::::::::::|||
QY 25 GlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyrPheGlyGlyGlu 44
   ||| ::::::::::::::::::::||| ::::::::::::::::::::||| ::::::::::::::::::::|||
Db 820 GGAAGATATAGTCACTCATCAATAAGACTGCATGCAGCTGAGCTGGTGGGAAGAGAG 879
   ||| ::::::::::::::::::::||| ::::::::::::::::::::||| ::::::::::::::::::::|||
QY 45 LeuLysGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIlePro 61
   ||| ::::::::::::::::::::||| ::::::::::::::::::::||| ::::::::::::::::::::|||
Db 880 CTGAACGGACAGACGAGCGGTTCGCCGATTAAGCTGTAAGTACTTCCA 930
   ||| ::::::::::::::::::::||| ::::::::::::::::::::||| ::::::::::::::::::::|||

RESULT 5
US-08-484-709-3
; Sequence 3, Application US/08/484,709
; Patent No. 5837844
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,709
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-192
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1392 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-484-709-3

Alignment Scores:
Pred. No.: 5.98e-09      Length: 1392
Score: 129.00           Matches: 22
Percent Similarity: 64.91% Conservative: 15
Best Local Similarity: 38.60% Mismatches: 20
Query Match: 38.74%      Indels: 0
DB: 2                   Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-08-484-709-3 (1-1392)
QY 5 TyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThrIleGlnPro 24
   ||| ::::::::::::::::::::||| ::::::::::::::::::::||| ::::::::::::::::::::|||
Db 760 TACTGCAAGATATATTTCATATGAGGACAGAAATGATGATTAATTGACAAATCAAGAA 819
   ||| ::::::::::::::::::::||| ::::::::::::::::::::||| ::::::::::::::::::::|||
QY 25 GlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyrPheGlyGlyGlu 44
   ||| ::::::::::::::::::::||| ::::::::::::::::::::||| ::::::::::::::::::::|||
Db 820 GGAAGATATAGTCACTCATCAATAAGACTGCATGCAGCTGAGCTGGTGGGAAGAGAG 879
   ||| ::::::::::::::::::::||| ::::::::::::::::::::||| ::::::::::::::::::::|||
QY 45 LeuLysGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIlePro 61
   ||| ::::::::::::::::::::||| ::::::::::::::::::::||| ::::::::::::::::::::|||
Db 880 CTGAACGGACAGACGAGCGGTTCGCCGATTAAGCTGTAAGTACTTCCA 930
   ||| ::::::::::::::::::::||| ::::::::::::::::::::||| ::::::::::::::::::::|||

RESULT 6
US-08-474-697-3
; Sequence 3, Application US/08/474,697
; Patent No. 6171800
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,697
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-193
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
```

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1392 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-474-697-3

Alignment Scores:

Pred. No.: 5.98e-09 Length: 1392
Score: 129.00 Matches: 22
Percent Similarity: 64.91% Conservative: 15
Best Local Similarity: 38.60% Mismatches: 20
Query Match: 38.74% Indels: 0
DB: 3 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-08-474-697-3 (1-1392)

QY 5 TyTyrAgaAlaLeuYrProPheGluSerArgSerHisAspGluIleThrIleGlnPro 24

Db 760 TACTGCAAGTAATATTTCCATATGAGCAGACAGATGATGATGATGACATCAAGAA 819

QY 25 GlyAspIleValMetValAspGluSerGlnThrGlyGlnProGlyTyrLeuGlyGlyGlu 44

Db 820 GGAGATATGACTCTCATATCAATAGAGACTGCATCGACGTAGGCTGGTGGAAAGAGAG 879

QY 45 LeuysgIlysthrGlyTyrPheProAlaAsnTyrAlaGluIlysIlePro 61

Db 880 CTGAACGGCAGACAGAGGCGTGTTCCTCCGATTACTTCGTGAAGTTACTTCCA 930

RESULT 7

US-08-475-894-1
Sequence 1, Application US/08475894
Patent No. 5641748

GENERAL INFORMATION:

APPLICANT: Yen-Ming Hsu
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,894

FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-191

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1659 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-475-894-1

Alignment Scores:
Pred. No.: 7.53e-09 Length: 1659

Score: 129.00 Matches: 22
Percent Similarity: 64.91% Conservative: 15
Best Local Similarity: 38.60% Mismatches: 20
Query Match: 38.74% Indels: 0
DB: 3 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-08-475-894-1 (1-1659)

QY 5 TyTyrAgaAlaLeuYrProPheGluSerArgSerHisAspGluIleThrIleGlnPro 24

Db 1027 TACTGCAAGTAATATTTCCATATGAGCAGACAGATGATGATGATGACATCAAGAA 1086

QY 25 GlyAspIleValMetValAspGluSerGlnThrGlyGlnProGlyTyrLeuGlyGlyGlu 44

Db 1087 GGAGATATGACTCTCATATCAATAGAGACTGCATCGACGTAGGCTGGTGGAAAGAGAG 1146

QY 45 LeuysgIlysthrGlyTyrPheProAlaAsnTyrAlaGluIlysIlePro 61

Db 1147 CTGAACGGCAGACAGAGGCGTGTTCCTCCGATTACTTCGTGAAGTTACTTCCA 1197

RESULT 8

US-08-484-710-1
Sequence 1, Application US/08484710
Patent No. 5656438

GENERAL INFORMATION:

APPLICANT: Yen-Ming Hsu
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,710
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1659 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-484-710-1

Alignment Scores:
Pred. No.: 7.53e-09 Length: 1659

Score: 129.00 Matches: 22
Percent Similarity: 64.91% Conservative: 15
Best Local Similarity: 38.60% Mismatches: 20
Query Match: 38.74% Indels: 0
DB: 3 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-08-484-710-1 (1-1659)

QY 5 TyTyrAgaAlaLeuYrProPheGluSerArgSerHisAspGluIleThrIleGlnPro 24

Db 1027 TACTGCAAGTAATATTTCCATATGAGCAGACAGATGATGATGATGACATCAAGAA 1086

```

QY      25  GIVASPILEVALMEVLAISPGLUSERGINHRRGLYGIUPROGILYTPPLEUGLIGLYGLU  44
           |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB      1087  GGAGATATATGCTACTCTCATCAATATAGACCTGCATCGACGTAGGCTGGTGGGAAGAGAG  1146
           |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
QY      45  LEUYSGLIYSTRHGLYTRPHEPROALAASNTYRALAGLUYSILEPRO  61
           |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB      1147  CTGAACGGCAGACGAGCGGTGTCTCCCGCATTAACCTCGGAAGTTACTTCCA  1197
           |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:

RESULT 9
US-08-484-709-1
; Sequence 1, Application US/08484709
; Patent No. 5837844
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,709
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-192
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1659 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-484-709-1

Alignment Scores:
Pred. No.: 7,53e-09 Length: 1659
Score: 129.00 Matches: 22
Percent Similarity: 64.91% Conservative: 15
Best Local Similarity: 38.60% Mismatches: 20
Query Match: 38.74% Indels: 0
DB: 2 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-08-484-709-1 (1-1659)
QY      5  TYTYRATGALALEUTYTRPHEGINSERARGSERHISAEPGLUEHTRILECLNPRO  24
           :::::|||||.....:|||||.....:|||||.....:|||||.....:
DB      1027  TACTCCAAAGTATATTTCCATATGAGCGACAGATGATGATGAATTGACATCAAGAA  1086
           :::::|||||.....:|||||.....:|||||.....:|||||.....:
QY      25  GIVASPILEVALMEVLAISPGLUSERGINHRRGLYGIUPROGILYTPPLEUGLIGLYGLU  44
           |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB      1087  GGAGATATATGCTACTCTCATCAATATAGACCTGCATCGACGTAGGCTGGTGGGAAGAGAG  1146
           |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
QY      45  LEUYSGLIYSTRHGLYTRPHEPROALAASNTYRALAGLUYSILEPRO  61
           |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB      1147  CTGAACGGCAGACGAGCGGTGTCTCCCGCATTAACCTCGGAAGTTACTTCCA  1197
           |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:

RESULT 10
US-08-474-697-1

```

```

; Sequence 1 Application US/08474697
; Patent No. 6171800
;
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,697
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 1:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1659 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
; US-08-474-697-1
;
; Alignment Scores:
; Pred. No.: 7.53e-09 Length: 1659
; Score: 129.00 Matches: 22
; Percent Similarity: 64.91% Conservative: 15
; Best Local Similarity: 38.60% Mismatches: 20
; Query Match: 38.74% Indels: 0
; Gaps: 3
;
; US-09-720-934-2_COPY_740_800 (1-61) x US-08-474-697-1 (1-1659)
;
Qy 5 TyrTyrAgtgAlaLeuTYrProPhgcluseTsrYgSerhTsAspguileEtrHleGlnPro 24
|||.....:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1027 TACGCAAGAGAAATATTTCCATATGAGGACACAGAAATGATGTAATGACATCAAAAGA 1086
25 GlysApIleValMetValAspGluSerGlnThrGlygluProGlyTYrPleuGlyGlyGlu 44
|||.....:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1087 GGAGATATAGCACTCATCATCAATAAAGACTGCATGCAGCGAGGCTGTGGGAAGAGAG 1146
45 LeuIysGlyStThrGlyTYrPhaProAlaAsnTYrAlaGluIysIlePro 61
Db 1147 CTGAACGGCAACGACGAGGCGTGTCTCCGATPACTTCGTGAAGTACTTCCA 1197
;
; RESULT 11
; US-08-671-354-1
; Sequence 1, Application US/08671354
; Patent No. 6423824
;
; GENERAL INFORMATION:
; APPLICANT: Hsu, Yen-Ming
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
;
;
;

```

1	COUNTRY: USA	
2	ZIP: 02110-2804	
3	COMPUTER READABLE FORM:	
4	MEDIUM TYPE: Diskette	
5	COMPUTER: IBM Compatible	
6	OPERATING SYSTEM: Windows 95	
7	SOFTWARE: FastSeq for Windows Version 2.0b	
8	CURRENT APPLICATION DATA:	
9	APPLICATION NUMBER: US/08/671.354	
10	FILING DATE: 27-JUN-1996	
11	PRIOR APPLICATION DATA:	
12	APPLICATION NUMBER: US 08/486,344	
13	FILING DATE: 07-JUN-1995	
14	APPLICATION NUMBER: US 08/484,709	
15	FILING DATE: 07-JUN-1995	
16	APPLICATION NUMBER: US 08/475,894	
17	FILING DATE: 07-JAN-1995	
18	APPLICATION NUMBER: US 08/475,710	
19	FILING DATE: 07-JAN-1995	
20	APPLICATION NUMBER: US 08/474,697	
21	FILING DATE: 07-JUN-1995	
22	ATTORNEY/AGENT INFORMATION:	
23	NAME: Myers, Louis	
24	REGISTRATION NUMBER: 35,965	
25	REFERENCE/DOCKET NUMBER: 10274/009005	
26	TELECOMMUNICATION INFORMATION:	
27	TELEPHONE: 617/542-5070	
28	TELEFAX: 617/542-8906	
29	TELEX: 200154	
30	INFORMATION FOR SEQ ID NO: 1:	
31	SEQUENCE CHARACTERISTICS:	
32	LENGTH: 3143 base pairs	
33	TYPE: nucleic acid	
34	STRANDEDNESS: single	
35	TOPOLOGY: linear	
36	MOLECULE TYPE: cDNA	
37	FEATURE:	
38	NAME/KEY: Coding Sequence	
39	LOCATION: 436...2412	
40	US-08-671-354-1	
41	Alignment Scores:	
42	Pred. No.: 1.75e-08	Length: 3143
43	Score: 129.00	Matches: 22
44	Percent Similarity: 64.91%	Conservative: 15
45	Best Local Similarity: 38.60%	Mismatches: 20
46	Query Match: 38.74%	Indels: 0
47	DB: 4	Gaps: 0
48	US-09-720-934-2_COPY_740_800 (1-61) * US-08-671-354-1 (1-3143)	
49	QY 5 TyrTyrArgAlaLeuTyrProPheGluSerArgSerHisaspGluIleThrIleGlnPro 24	
50	Db 1228 TACGCAAGAAATTAATTCCTCATATAGAGGCACAGAAATGATGATGAATTCACATCAAAAGAA 1287	
51	QY 25 GlVAsplIeValMetValAspGluSerGlnThrGlyGluProGlyTyrPLeuGlyGlyGlu 44	
52	Db 1288 GGAGATTAAGTCACTCTTCATCATATAAGACCTGCATCCAGCGTAGGCTGGTGGAAAGAGAG 1347	
53	QY 45 LeuIeGlyIeSthGlyTyrPheProAlaAsnTyrAlaGluIlePro 61	
54	Db 1348 CTGAACGCGACAGAGCGCTCTCCCGATTAACCTTGGAAGTAACTTCCA 1398	
55	RESULT 12	
56	US-08-306-691B-48	
57	/ Sequence 48, Application US/08306691B	
58	/ Patent No. 5734039	
59	/ GENERAL INFORMATION:	
60	/ APPLICANT: Calabretta, Bruno	
61	/ APPLICANT: Skorski, Tomasz	
62	/ TITLE OF INVENTION: ANTISENSE	
63	/ TITLE OF INVENTION: OligonucleotideTTDBS TARGETING COOPERATING ONCOGENES	
64	/ NUMBER OF SEQUENCES: 55	

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Seidel, Gonda, Lavorjona & Monaco, P.C.
STREET: Two Penn Center, Suite 1800
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19102

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 KB
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,691B
FILING DATE: September 15, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5734039e
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 2757 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-306-691B-48

Alignment Scores:
Pred. No.: 1.73e-08 Length: 2757
Score: 128.50 Matches: 22
Percent Similarity: 67.92% Conservative: 14
Best Local Similarity: 41.51% Mismatches: 16
Query Match: 38.59% Indels: 1
DB: 1 Gaps: 1

US-09-720-934-2_COPY 740_800 (1-61) x US-08-306-691B-48 (1-2757)
QY 7 ArgAlaLeuTyrProPheGluSerArgSerHisAspGluLeuThrIleGlnProGlyAsp 26
Db 2328 AAAGCCCGCATATGACTTCTGGGCCCGTCAAGCGTTCAGAGCTGCGTCAAGAGAGGTGAC 2387
QY 27 IleValMetValAspGluSerGlnThrGlyGluProGlyTyrPleuGlyGlyGluLeuLys 46
Db 2388 ATCTCAAGATC--CTTAACAGAAAGAGGACAGCAAGAGCTGTGGCGAAGGAGAGATCTAT 2444
QY 47 GlyIythrGlyTyrPheProAlaAsnTyrAlaGluLys 59
Db 2445 GGCCGGGTGGCTGTGTTCCCTGCCCACTTACGTGAGGAA 2483

RESULT 13
PCT-US93-06251-79
Sequence 79, Application PC/RUS9306251
GENERAL INFORMATION:
APPLICANT: Wickstrom, Eric and Rife, Jason P.
TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06251
FILING DATE: 19930630
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: DIGILLO, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 2757 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-79

Alignment Scores:
Pred. No.: 1.73e-08 Length: 2757
Score: 128.50 Matches: 22
Percent Similarity: 67.92% Conservative: 14
Best Local Similarity: 41.51% Mismatches: 16
Query Match: 38.59% Indels: 1
DB: 5 Gaps: 1

US-09-720-934-2_COPY_740_800 (1-61) x PCT-US93-06251-79 (1-2757)

QY 7 ArgAlaLeuTyPrpHeGluSerHsAspGluIleThrIleGlnProGlyAsp 26

Db 2328 AAAGCCGCTAGCTTCTGCGCCGTCACCGTTCAGAGCTGTCGCTCAAGAGGGGTAC 2387

QY 27 lIeValMetValAspGluSerGlnThrGlyGluProGlyTyrPleuGlyGlyIuLeuLys 46

Db 2388 ATCATCAAGATC---CTTAACAAGAGGACAGCAAGGCTGTCGCGAGGGGAGATCTAT 2444

QY 47 GlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLys 59

Db 2445 GGCCGGGTTGGCTGTTCCCTGCCACTAGTGAAGAA 2483

RESULT 14
US-07-646-537B-1
Sequence 1, Application US/07646537B

PATENT No. 5348664
GENERAL INFORMATION:
APPLICANT: Barbacid, Mariano
TITLE OF INVENTION: Yav Proto-Oncogene Protein
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bristol-Myers Squibb Company
STREET: P.O. Box 4000
CITY: Princeton
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08543-4000
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/646,537B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gaul, Timothy J.

REGISTRATION NUMBER: 33,111
REFERENCE/DOCKET NUMBER: DC10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 921-5901
TELEFAX: (609) 921-4526
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2793 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 14..2545
US-07-646-537B-1

Alignment Scores:
Pred. No.: 6.55e-08 Length: 2793
Score: 124.50 Matches: 21
Percent Similarity: 67.92% Conservative: 15
Best Local Similarity: 39.62% Mismatches: 16
Query Match: 37.39% Indels: 1
DB: 1 Gaps: 1

US-09-720-934-2_COPY_740_800 (1-61) x US-07-646-537B-1 (1-2793)

QY 7 ArgAlaLeuTyPrpHeGluSerHsAspGluIleThrIleGlnProGlyAsp 26

Db 2372 AAAGCCGCTAGCACTTCTGTCGCCGACAGAGTGGAGTGTCCCTTAAGAGGGGTGAT 2431

QY 27 lIeValMetValAspGluSerGlnThrGlyGluProGlyTyrPleuGlyGlyIuLeuLys 46

Db 2432 ATCATCAAGATC---CTCAATAAGAGGACAGCAAGGCTGTCGCGAGGAGATCTAC 2488

QY 47 GlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLys 59

Db 2489 GGCCGGATCGGCTGTTCCCTTCAACTAGTGAAGAA 2527

RESULT 15
US-08-630-915A-7
Sequence 7, Application US/08630915A

PATENT No. 6309820
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLES, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penite & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mierock, S. Leslie
REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1277 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-630-915A-7

Alignment Scores:
Pred. No.: 2.74e-07
Score: 117.00
Percent Similarity: 57.81%
Best Local Similarity: 43.75%
Query Match: 35.14%
DB: 4

Length: 1277
Matches: 28
Conservative: 9
Mismatches: 19
Indels: 8
Gaps: 3

US-09-720-934-2_COPY_740_800 (1-61) * US-08-630-915A-7 (1-1277)

QY	2	LySValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThr	21
DB	155	CAAGTTAAAGCTTCAGAGCTCTATATACATTGAACTCCAGACTCCAGATGAATTATAC	214
QY	22	IlleGlnProGlyAspIleValMetVal--AspGluSerGlnThrGlyGlnProGlyTyr	40
DB	215	TTTGAAAGAGGAGCATTTACTACATCATGACATGACATGATGATACC-----AGCTGG	265
QY	41	LeuGlyGlyGlyGluLeuGlyGlyLeuThrGlyTyrPheProAlaAsnTyr-----	56
DB	266	TGGAAGAGGACATGCAAGGCGCAAGACGACTGATCCGAGCAACTATGTGGCTGAGCAG	325
QY	57	AlaGluLysIle	60
DB	326	GCGAATCCATT	337

Search completed: July 1, 2004, 19:53:46
Job time : 47.7143 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:09:20 ; Search time 187.143 Seconds
(without alignments)
1572.471 Million cell updates/sec

Title: US-09-720-934-2_COPY_740_800
Perfect score: 333
Sequence: 1 VKVYVYRALYPBESRSHDEL.....GGELKCKTGMFPANNAEKIP 61

Scoring table:
BLOSUM62 Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3163042 segs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xip
-O/cgpn2_1/USPTO.spool.p/US09720934/runat.30062004.064541.13481/app.query.fasta_1.1386
-DB=Published Applications NA -QFMT=fastp -SUFFIX=rnpb -MINMATCH=0-1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62
-TRANS=human40.csl -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09720934 @CGN 1 1 1500 @runat.30062004.064541.13481
-NCPU=6 -ICPU=3 -NO MAP -LARGESQUERY -NEG SCORE=0 -MAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YCAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: Published Applications NA:*
2: /cgpn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
3: /cgpn2_6/ptodata/1/pubpna/PTC_NEW_PUB.seq:*
4: /cgpn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
5: /cgpn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgpn2_6/ptodata/1/pubpna/PTCUS_PUBCOMB.seq:*
7: /cgpn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgpn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgpn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgpn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgpn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgpn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq:*
13: /cgpn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
14: /cgpn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgpn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgpn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgpn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgpn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgpn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	333	100.0	3319	11	US-09-764-875-88
2	333	100.0	3466	16	US-10-158-057-33
3	216	64.9	3746	11	US-09-764-875-176
4	216	64.9	4210	9	US-09-764-868-125
5	216	64.9	5828	13	US-10-398-885A-15
6	215	64.6	2673	9	US-09-879-957-193
7	182	54.7	270	9	US-09-864-761-17127
8	182	54.7	286	9	US-09-864-761-26948
9	182	54.7	297	9	US-09-864-761-30453
10	182	54.7	475	9	US-09-864-761-311
11	182	54.7	475	9	US-09-864-761-13884
12	182	54.7	480	9	US-09-864-761-10314
13	140	42.0	970	9	US-09-833-381-1294
14	140	42.0	2539	16	US-10-104-047-1655
15	139	41.7	581	10	US-09-918-995-14907
16	138	41.4	1558	9	US-09-939-825-6
17	138	41.4	1558	9	US-09-764-868-124
18	138	41.4	1558	11	US-09-764-875-309
19	129	38.7	356	13	US-10-085-783A-37196
20	129	38.7	386	16	US-10-242-535A-37196
21	129	38.7	400	13	US-10-085-783A-22639
22	129	38.7	400	16	US-10-242-535A-22639
23	129	38.7	1215	17	US-10-648-593-61
24	129	38.7	3143	14	US-10-144-621-1
25	129	38.7	3348	13	US-10-342-887-187
26	129	38.7	3348	13	US-10-172-118-187
27	128.5	38.6	518	9	US-09-796-692-5870
28	128.5	38.6	518	15	US-10-040-862-5870
29	128.5	38.6	518	16	US-10-057-475B-5870
30	128.5	38.6	518	16	US-10-154-884B-5870
31	128.5	38.6	2757	10	US-09-960-706-1059
32	124	37.2	337	9	US-09-924-035A-843
33	124	37.2	337	13	US-09-770-423-865
34	124	37.2	564	17	US-10-029-386-15
35	123	36.9	1326	15	US-10-437-963-53573
36	122	36.6	474	9	US-09-864-761-31955
37	122	36.6	474	9	US-09-864-761-15441
38	122	36.6	1890	13	US-10-424-599-12461
39	121	36.3	1592	17	US-10-437-963-60380
40	118	35.4	1702	13	US-10-424-599-19639
41	117.5	35.3	3027	13	US-10-087-192-1871
42	117.5	35.3	3027	13	US-10-042-865-15
43	117.5	35.3	3141	16	US-10-258-106-27
44	117.5	35.3	3538	13	US-10-362-892-36
45	117.5	35.3	3538	16	US-10-288-798-36

ALIGNMENTS

RESULT 1
US-09-764-875-88 Application US/09764875
Sequence 88, Application US/09764875
Publication No. US20040018969A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P202
CURRENT APPLICATION NUMBER: US/09/764,875
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PAM or file wrapper
NUMBER OF SEQ ID NOS: 1249
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 88
LENGTH: 3319
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-875-88

Alignment Scores: 2.61e-39 Length: 3319
Pred. No.: 333.00 Matches: 61
Score: 333.00
Percent Similarity: 100.00% Conservative: 0

```
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 11
```

```
Mismatches: 0
Indels:      0
Gaps:        0
```

US-09-720-934-2_COPY_740_800 (1-61) X US-09-764-875-88 (1-3319)

QY 1 VallysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIle 20

DB 2388 G1AAAG1GG1GTAT1ACCGGCACTGTACCCCTTGAATCCAGAAGCCATGATGAATC 2447

21 nrlieglnprogl yaspil ealmetvalasp guserglntrhrg lyluproglytrp 40

2448 ACATACCAACGAGACATAGTCAATGATGAAAGCCAACTGGAGAACCCGGCTGG 2507

11 ucua ygi ygi aueuu ybgi yds imi oi i i pferioai aabim y iai aig iul y si e 80

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818

Db 2568 CCA 2570

RESULT 2
US-10-158-057-33

OTHER INFORMATION: n equals a,t,b, or c
US-10-158-057-33

Alignment Scores:

```
Pred. No.: 2.76e-3
Score: 333.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 16
```

Length:	3466
Matches:	61
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-720-934-2_COPY_740_800 (1-61) X US-10-158-057-33 (1-3466)

QY 1 VallysVal1TyrrArgAlaLeuTyrrProphGluSerArgSerHisAspGluIle 20

Db 2506 GTAAAGTGGTGTATTACCGGGCACTGTACCCCTTTGAATCCAGAAGCCATGATGAAATC 2565

21 ThrIleGlnProGlyAspIleValMetValAspGlnUserGlnThrGlyGlnProGlyTyr 40

Db 2566 ACTATCCAGCCAGAGACATAGTCATGGTGGATGAAGCCAACTGGAGAACCCGGCTGG 2625

41 LeuGlyGlyGluLeuLysGlyLysIrrpneproalaasntYrAaglulysile 60

20 2020 C1GGGCGAGAA1ATTTCGGATGACACGG1GG11CC1GCAAC1A1GCGAGAGAAATC 2685

$\frac{1}{2}$

Db 2686 CCA 2688

RESULT 3
ITC-09-764-87E-176

US-09-1764-873-176
; Sequence 176, Application US/09764875
; Publication No. US20040018969A1

US-09-764-875-176

Alignment Scores

Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

9.87e-22	Length:	3746
216.00	Matches:	36
74.58%	Conservative:	8
61.02%	Mismatches:	15
64.86%	Indels:	0
11	Gaps:	0

US-09-720-934-2_COPY_740_800 (1-61) x US-09-764-875-176 (1-3746)

3 ValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThrIle 22

Db 154 TTGGTGAATTATAGAGCATTATACCCCTTTGAAGCAAGAACCATGATGAGATGAGTTT 213

23 GlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTrpLeuGly 42

Db 214 AATTCTGAGATATAATTCAAGTTGATGAATAAACCGTAGGAGAACCTGGTTGGCTTAT 273

43 GLYGLULELysGLYLSTInGLYIrpheproAlaasntYrAlaGLULysIlepro 61

2/4 GGAGGIIICAGGAAIIIGGCGGIIICCAIGCAATTAIGTAGAAAAAATGCCA 330

US-09-76

; Sequence 125, Application US/09764868
; Patent No. US2002016871A1

ORGANISM: Homo sapiens
US-09-764-868-125

Alignment Scores:

```
Pred. No.: 1.14e-21
Score: 216.00
Percent Similarity: 74.58%
Best Local Similarity: 61.02%
Query Match: 64.86%
DB: 9
```

Length:	4210
Matches:	36
Conservative:	8
Mismatches:	15
Indels:	0
Gaps:	0

US-09-720-934-2 COPY 740 800 (1-61) x US-09-764-868-125 (1-4210)

QY 3 ValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThrIle 22


```
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 32
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.66
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
; OTHER INFORMATION: NT HIT: AF114488.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q15811, EVALUE 2.00e-45
; OTHER INFORMATION: EST_HUMAN HIT: AA773263.1, EVALUE 1.00e-112
US-09-864-761-26948

Alignment Scores:
Pred. No.: 4,93e-18 Length: 286
Score: 182.00 Matches: 32
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 54.65% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) * US-09-864-761-26948 (1-286)
QY 30 ValAspGluSerGlnThrGlyGluProGlyTyrIleuGlyGlyGluLeuGlyGlyThr 49
DB 1 GTGGATGAAAGCCAACTGGAGAACCCGGCTGGCTTGAGAGAGATTAAAGGAAAGACA 60
QY 50 GlyTyrPheProAlaAsnTyrAlaGluTyrIlePro 61
DB 61 GGGTGGTTCCTCCGCAAACTATGACAGAGAAATCCCA 96

RESULT 9
US-09-864-761-30453
; Sequence 30453, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 30453
LENGTH: 297
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000117.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.95
; OTHER INFORMATION: NT HIT: AF114487.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: AA773263.1, EVALUE 1.00e-112
; OTHER INFORMATION: SWISSPROT HIT: Q15811, EVALUE 2.00e-45
US-09-864-761-30453

Alignment Scores:
Pred. No.: 5.17e-18 Length: 297
Score: 182.00 Matches: 32
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 54.65% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) * US-09-864-761-30453 (1-297)
QY 30 ValAspGluSerGlnThrGlyGluProGlyTyrIleuGlyGlyGluLeuGlyGlyThr 49
DB 1 GTGGATGAAAGCCAACTGGAGAACCCGGCTGGCTTGAGAGAGATTAAAGGAAAGACA 60
QY 50 GlyTyrPheProAlaAsnTyrAlaGluTyrIlePro 61
DB 61 GGGTGGTTCCTCCGCAAACTATGACAGAGAAATCCCA 96

RESULT 10
US-09-864-761-311
; Sequence 311, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
```

```
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 311
/ LENGTH: 475
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AF000193.1
/ OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.2
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.2
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
/ US-09-864-761-311

Alignment Scores:
Pred. No.: 9,27e-18 Length: 475
Score: 182.00 Matches: 32
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 54.65% Indels: 0
Gaps: 0
DB: 9

US-09-720-934-2_COPY_740_800 (1-61) x US-09-864-761-311 (1-475)
QY 30 ValaspguserglnthrglyglnpProglyTrrpleuglygyluleuLyseGlyLysThr 49
Db 275 GTGATGAAAGCCAAACTGGAAGAACCCGCTGCTGAGAGAAATTAAAAAGAAAGACA 334
QY 50 G1YrrpPheProAlaAsnTYrAlaGluLysIlePro 61
Db 335 GGGTGTTCCCTGCACAACTATGACAGAGAAATCCCA 370

RESULT 11
US-09-864-761-13884
/ Sequence 13884, Application US/09864761
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
/ FILE REFERENCE: Acomica-x-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
```

```
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 13884
/ LENGTH: 475
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AP000117.1
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.7
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
/ OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.76
/ OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.3
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.95
/ US-09-864-761-13884

Alignment Scores:
Pred. No.: 9,27e-18 Length: 475
Score: 182.00 Matches: 32
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 54.65% Indels: 0
Gaps: 0
DB: 9

US-09-720-934-2_COPY_740_800 (1-61) x US-09-864-761-13884 (1-475)
QY 30 ValaspguserglnthrglyglnpProglyTrrpleuglygyluleuLyseGlyLysThr 49
Db 275 GTGATGAAAGCCAAACTGGAAGAACCCGCTGCTGAGAGAAATTAAAAAGAAAGACA 334
QY 50 G1YrrpPheProAlaAsnTYrAlaGluLysIlePro 61
Db 335 GGGTGTTCCCTGCACAACTATGACAGAGAAATCCCA 370
```

```

RESULT 12
US-09-864-761-10314
; Sequence 10314, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharton G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 10314
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF000311.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.87
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 32
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.66
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
US-09-864-761-10314

Alignment Scores:
Pred. No.: 9,39e-18 Length: 480
Score: 182.00 Matches: 32
Percent Similarity: 100.00% Conservative: 0

```

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 54.65% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x US-09-864-761-10314 (1-480)

QY 30 ValAspGluSerGlyThrGlyGluProGlyTyrPleuGlyGlyGluLeuGlyLysThr 49
DB 243 GGGAGTGAAGGCAAGCTGGAGAACCCGGCTGGCTGGAGAGATTAAAGGAAAGACA 302

QY 50 GlyTyrPheProAlaAsnTyrAlaGlyLysLeuPro 61
DB 303 GGGTGGTTCCTGCCTGCAAACTATGCGAGGAAATCCCA 338

RESULT 13
US-09-833-381-1294/c
; Sequence 1294, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1294
; LENGTH: 970
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1294

Alignment Scores:
Pred. No.: 4.36e-11 Length: 970
Score: 140.00 Matches: 26
Percent Similarity: 67.27% Conservative: 11
Best Local Similarity: 47.27% Mismatches: 16
Query Match: 42.04% Indels: 2
DB: 9 Gaps: 1

US-09-720-934-2_COPY_740_800 (1-61) x US-09-833-381-1294 (1-970)

QY 7 ArgAlaLeuTyrProPheGluSerArgSerHisAspGluLeuThrIleGlnProGlyAsp 26
DB 376 AAGCGCTTGATTCCTTTTCAGCCAGCAAGATGATGATTTGGAAAGGGTGAC 317

QY 27 IleValMetValAspGluSerGlnThrGlyGluProGlyTyrPleuGlyGlyGluLeuLys 46
DB 316 ATTGTGATTATACACGAGAAAAA-----GAGGAAAGATGGTGGTTGATCTTTGAAT 263

QY 47 GlyLysThrGlyTyrPheProAlaAsnTyrAlaGlyLysLeuPro 61
DB 262 GGGAAAAAGGCCATTTCCTCGCGCTTATGCGAGAGATTACT 218

RESULT 14
US-10-104-047-1655
; Sequence 1655, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392A1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1655
; LENGTH: 2539
; TYPE: DNA

```

ORGANISM: Homo sapiens
US-10-104-047-1655

Search completed: July 1, 2004, 20:15:55
Job time : 193.143 secs

Alignment Scores:

Pred. No.:	1.44e-10	Length:	2539
Score:	140.00	Matches:	26
Percent Similarity:	67.27%	Conservative:	11
Best Local Similarity:	47.27%	Mismatches:	16
Query Match:	42.04%	Indels:	2
DB:	16	Gaps:	1

US-09-720-934-2_COPY_740_800 (1-61) X US-10-104-047-1655 (1-2539)

QY 7 ArgAlaLeuTYrProPheGluSerArgSerHisAspGluIleThrIleGlnProGlyAsp 26
Db 2257 AAGGCTTGATTTCTTTCAAGCAGGACAGATGATGATGATTTGGAAAAAGGTGAC 2316

QY 27 IleValMetValAspGluSerGlnThrGlyGluProGlyTyrPleuGlyGlyGluLeuLys 46
Db 2317 ATTGTGATTATACACGAGAAAAA-----GAGAGAGATGTGTGTTGATCTTGAAT 2370

QY 47 GlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIlePro 61
Db 2371 GGGAAAAAGGCCATTTCCTGCGCCTTATGTGAGAGATTACT 2415

RESULT 15

US-09-918-995-14907
Sequence 14907, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995

PRIOR FILING DATE: 2001-07-30
PRIORITY APPLICATION NUMBER: US/09/235,076

NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14907

LENGTH: 581
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(581)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-14907

Alignment Scores:

Pred. No.:	3.25e-11	Length:	581
Score:	139.00	Matches:	26
Percent Similarity:	67.27%	Conservative:	11
Best Local Similarity:	47.27%	Mismatches:	16
Query Match:	41.74%	Indels:	2
DB:	10	Gaps:	1

US-09-720-934-2_COPY_740_800 (1-61) X US-09-918-995-14907 (1-581)

QY 7 ArgAlaLeuTYrProPheGluSerArgSerHisAspGluIleThrIleGlnProGlyAsp 26
Db 307 AAGGCTTGATTTCTTTCAAGCAGGACAGATGATGATGATTTGGAAAAAGGTGAC 366

QY 27 IleValMetValAspGluSerGlnThrGlyGluProGlyTyrPleuGlyGlyGluLeuLys 46
Db 367 ATTGTGATTATACACGAGAAAAA-----GAGAGAGATGTGTGTTGATCTTGAAT 420

QY 47 GlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIlePro 61
Db 421 GGGAAAAAGGCCATTTCCTGCGCCTTATGTGAGAGATTACT 465

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:01:26 ; Search time 1835 Seconds
(without alignments)
992.694 Million cell updates/sec

Title: US-09-720-934-2_COPY_740_800
Perfect score: 333
Sequence: 1 VAVVYIRALYPRESHSDI.....GGLKCKTGWFPANAEKIP 61

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xip
-O=/cgm2.1/USPTO.spool.p/US09703934/runat.30062004.064540.13442/app.query.fasta_1.1386
-DB=EST -OPMT=fastap -SUFFIX=xt -MINMATCH=0.1 -ICOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US09703934 @CGM 1.1 12421 @runat.30062004.064540.13442 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MAVN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em.estba:*
2: em.esthm:*
3: em.esthm:*
4: em.estmu:*
5: em.estov:*
6: em.estpl:*
7: em.estro:*
8: em.htc:*
9: gb.est1:*
10: gb.est2:*
11: gb.htc:*
12: gb.est3:*
13: gb.est4:*
14: gb.est5:*
15: em.estfun:*
16: em.estom:*
17: em.gss.hum:*
18: em.gss.hiv:*
19: em.gss.pln:*
20: em.gss.vtl:*
21: em.gss.fun:*
22: em.gss.man:*
23: em.gss.mus:*
24: em.gss.pro:*
25: em.gss.rtd:*
26: em.gss.phg:*
27: em.gss.vtl:*
28: gb.gss1:*

29: gb.gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	333	100.0	436	9	AL121153	AL121153 DKFZp762K05
2	333	100.0	488	9	AA773823	AA773823 at6d07.r
3	333	100.0	489	9	AA773263	AA773263 at62e11.r
4	333	100.0	527	10	BB757493	BB757493 BB757493
5	333	100.0	573	14	CA874698	CA874698 K0934C04
6	333	100.0	650	10	BB656585	BB656585 BB656585
7	333	100.0	657	10	AW950164	AW950164 EST362129
8	333	100.0	698	14	CF737796	CF737796 UI-M-HD0
9	333	100.0	5385	11	BC062938	BC062938 Mus. muscu
10	329	98.8	608	14	CA511580	CA511580 UI-R-FU0
11	320	96.1	703	14	CB527154	CB527154 UI-M-FY0
12	319	95.8	673	14	CF735070	CF735070 UI-M-HB0
13	317	95.2	443	9	AA492783	AA492783 v17307.r
14	302	90.7	600	14	CF738758	CF738758 UI-M-HD0
15	301	90.4	551	9	AV669700	AV669700 AV669700
16	293	88.0	481	9	AA212646	AA212646 m80c01.r
17	288	86.5	500	14	CD585892	CD585892 RK033A2H0
18	285	85.6	729	14	CF948685	CF948685 UI-M-HJ0
19	281	84.4	688	14	CF729086	CF729086 UI-M-HD0
20	277	83.2	456	9	AA313515	AA313515 EST185546
21	267	80.2	495	14	CA872596	CA872596 K091C006
22	244	73.3	148	9	AA670080	AA670080 at79a03.r
23	239	71.8	337	14	CD586870	CD586870 RK038A3B0
24	237	71.2	201	14	CF906343	CF906343 A0446A03
25	230	69.1	426	12	BC992086	BC992086 NR2-HT116
26	219	65.8	625	13	BU245991	BU245991 603290289
27	219	65.8	1133	13	BU144332	BU144332 603290289
28	217	65.2	833	12	BI872589	BI872589 UI-HF-BN0
29	216	64.9	653	14	CF136055	CF136055 UI-M-HD0
30	216	64.9	700	13	BK481721	BK481721 DKFZp686G
31	216	64.9	727	14	CF539609	CF539609 UI-M-GV0
32	216	64.9	797	12	BG862028	BG862028 602795041
33	216	64.9	951	12	BM465259	BM465259 ACENCOUPT
34	214	64.3	555	10	AA822841	AA822841 UG20F04.Y
35	201	60.4	774	13	BU316793	BU316793 603855209
36	198	59.5	384	9	AJ282752	AJ282752 4A3A-ABB
37	198	59.5	681	12	BM596246	BM596246 170006874
38	183	55.0	699	13	BM216116	BM216116 BM216116
39	182	54.7	587	9	AV595297	AV595297 AV595297
40	179	53.8	144	14	X94499	X94499 RNSAP13M.RA
41	178	53.5	527	14	CF729468	CF729468 UI-M-HD0
42	176	52.9	1059	29	CNS057EX	AL132456 Tetradon
43	176	52.9	1072	29	CNS02AR2	AL188903 Tetradon
44	167	50.2	649	12	BJ103848	BJ103848 BJ103848
45	164	49.2	630	13	BQ179493	BQ179493 UI-M-EW0

ALIGNMENTS

RESULT 1
LOCUS AL121153 436 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp762K055_r1 762 (synonym: lme12) Homo sapiens cDNA clone
ACCESSION AL121153
VERSION AL121153.1 GI:5927154
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 436)

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 489)
Hillier, L., Allen, M., Bowler, L., Dubuque, T., Giesel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Maita, M., Martin, J., Moore, B., Scheinberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNM; contact the IMAGE Consortium (info@image.lnml.gov) for further information.
Possible reversed clone: similarity on wrong strand
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 445.

FEATURES
Location/Qualifiers
1..489
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1048556"
/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab_host="DH10B"
/clone_lib="Soares NhhMPU_S1"
/note="Organ: mixed (see below); Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbhM, pregnant uterus NbhPU, and fetal heart NbhH19M) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1 M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

ORIGIN
Alignment Scores:
Pred. No.: 1,24e-31 Length: 489
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AA773263 (1-489)

QY 1 VallysyValValTYTYTYArgAlaLeuTYrProPhgGUsrArgSerHisAspGluIle 20
Db 175 GTAAAGGGGTGATTTACCGGGCACTGATCCCTTGATTCACAGAGCATGATGAAATC 234
QY 21 ThrlleGlnProGlyAspIleValMetValAspGluSerGlnThnGlyGluProGlyTTrp 40
Db 235 ACTATCCAGCCAGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 294
QY 41 LeuGlyGlyGluLeuLeuGlyGlyThrGlyTTrpPheProAlaSerTYAlaGluIle 60
Db 295 CTTGAGGAGAAATTAAAGAAAGACAGGCTGCTTCCCTGCAAACTATGAGAGAAATC 354
QY 61 Pro 61
Db 355 CCA 357

RESULT 4

BB757493
LOCUS 527 bp mRNA linear EST 16-OCT-2001
DEFINITION BB757493 RIKEN full-length enriched, melanocyte Mus musculus cDNA
Clone G270085L11 3', mRNA sequence.
ACCESSION BB757493
VERSION BB757493.1 GI:16185846
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 527)
Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Nutsaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Watanabe, A., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

TITLE Unpublished (2001)
COMMENT Contact: Yoshinide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10): 1617-1630 (2000)
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsuhiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.

FEATURES
Location/Qualifiers
1..527
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G270085L11"
/cell_type="melanocyte"
/clone_lib="RIKEN full-length enriched, melanocyte"

ORIGIN
Alignment Scores:
Pred. No.: 1,36e-31 Length: 527
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x BB757493 (1-527)

Matsunura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Komuro, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Kondo, S., Shinagawa, A., Salto, T., Kiyosawa, H., Yamataka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y. Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.

FEATURES

Location/Qualifiers

1..650
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="D130043E15"
/tissue="spinal ganglion"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 12 days embryo spinal ganglion"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGCGCGCCGCACTGAGTTTCTTTTCTTTTCTT 3']. cDNA was prepared by using triazole thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGTTCTCGAGTTAATTAATTAATTCCTCCCTCCCTCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Alignment Scores:

Pred. No.: 1..78e-31 Length: 650
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x BB656585 (1-650)

QY 1 VallysvaVallyTYTtYrAgaLaLeuTYrProPhgIuSerArGserHIsaSpGluIle 20
DB 8 GTAAAGTGTATTAACGAGCGCTGTACCCCTTGAATCCAGAACTCAGATGATC 67
QY 21 ThriIeGlnProGlnAspIleValMetValAspGlnSerGlnThrGlyGluProGlyTyr 40
DB 68 ACCATCCAGCCAGAGATATAGTCATGTGATGAAAGCCAGACTGAGAGCCAGATGG 127
QY 41 LeuGlyGlyGluLeuLeuGlyGlyTyrPheProAlaAsnTYrAlaGluIle 60
DB 128 CTGGAGAGAGCTGAAGGAGAGACGGATGTTCCCTCCAACTATGAGAAAGATT 187
QY 61 Pro 61
DB 188 CCA 190
RESULT 7

AW950164 657 bp mRNA linear EST 01-JUN-2000
LOCUS
DEFINITION EST362129 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.
ACCESSION AW950164
VERSION AW950164.1 GI:8139701
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 657)
Hedge, P., Qi, R., Abernethy, K., Dharap, S., Gaepard, R., Gay, C., Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and Quackenbush, J.

TITLE Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 10

FEATURES

Location/Qualifiers

1..657
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGE"
/note="Vector: pBluescriptSkm"

ORIGIN

Alignment Scores:

Pred. No.: 1..8e-31 Length: 657
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AW950164 (1-657)

QY 1 VallysvaVallyTYTtYrAgaLaLeuTYrProPhgIuSerArGserHIsaSpGluIle 20
DB 322 GTAAAGTGTATTAACGAGCGCTGTACCCCTTGAATCCAGAACTGATGAATC 381
QY 21 ThriIeGlnProGlnAspIleValMetValAspGlnSerGlnThrGlyGluProGlyTyr 40
DB 382 ACTATCCAGCCAGAGACATAGTCATGTGATGAAAGCCAACTGAGAAACCGGCTGG 441
QY 41 LeuGlyGlyGluLeuLeuGlyGlyTyrPheProAlaAsnTYrAlaGluIle 60
DB 442 CTGGAGAGAGATTAAAGGAGAGACAGGATGTTCCCTCCAACTATGAGAGAAATC 501
QY 61 Pro 61
DB 502 CCA 504
RESULT 8
CF737796 698 bp mRNA linear EST 10-OCT-2003
LOCUS
DEFINITION UT-M-HD0-ctc-i-11-0-UT-r1 NIH_BMAP_HD0 Mus musculus cDNA clone
ACCESSION CF737796
VERSION CF737796.1 GI:37634133
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 698)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. James Lin University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouse1.html>
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES
 Source
 Location/Qualifiers
 1. 698
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:36611818"
 /tissue_type="whole eye"
 /dev_stage="embryo 12.5,13.5,14.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP HD0"
 /note="Organ: Eye; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction. Ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TTATTGAAGT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

ORIGIN
 Alignment Scores:
 Pred. No.: 1,95e-31 Length: 698
 Score: 333.00 Matches: 61
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) * CFT37796 (1-698)

QY 1 VallystValValYTYTATGAAAlaLeuTYrProPhcGluSerArgSerHisAspGluIle 20
 Db 240 GTAAAGGGATATTAACCGAGGCGCTGACCCCTTGATCCAGAAAGTCAGAGAGATC 299
 QY 21 ThrllleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr 40
 Db 300 ACCATCCAGCCAGAGATATGTCATGCTGATGAAGCCAGATCTGAGAGCCAGATGG 359
 QY 41 LeuGlyGlyGluLeuLeuysGlyLysThrGlyTyrPheProAlaAsnTYrAlaGluLysIle 60
 Db 360 CTTCGAGAGAGAGTGAAGAGGAGACGCGGATGTTCCCTCGCAAACTATGACGAAAGATT 419
 QY 61 Pro 61
 Db 420 CCA 422

RESULT 9
 BC062938 \$385 bp mRNA linear HTC 11-DEC-2003
 LOCUS Mus musculus interectin (SH3 domain protein 1A), mRNA (cDNA clone IMAGE:6839463), containing frame-shift errors.
 ACCESSION BC062938

VERSION BC062938.1 GI:38566052
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 5385)

AUTHORS
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldi, M.P., Casavant, T.L., Schaefer, T.E., Brownstein, M.J., Udén, T.B., Tschirpki, S., Carninci, P., Prange, C., Kaha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullanly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gamaralline, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Bulky, S.W., Vallalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Scherchenko, Y., Bonfield, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 5385)
 Strausberg, R.

TITLE Direct Submission
JOURNAL Submitted (26-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 CDNA Library Preparation: M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLT)
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
 Web site: <http://genome.uiowa.edu>
 Contact: bento-soares@uiowa.edu, tom-casavant@uiowa.edu
 Bonaldi, M.F., Akabogu, I., Bair, T., Bair, U., Crouch, K., Davis, A., Fishler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Snit, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

ORIGIN
 Alignment Scores:
 Pred. No.: 2.63e-30 Length: 5385
 Score: 333.00 Matches: 61
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

FEATURES
 source
 location/Qualifiers
 1. 5385
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6839463"
 /tissue_type="Brain"
 /clone_lib="NIH BMAP_GH0"
 /lab_host="DH10B"
 /note="Vector: pYX-Asc"

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLT at: <http://image.llnl.gov>
 Series: Plate: Row: Column: 0
 This clone has the following problem: frame shifted.

Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0
US-09-720-934-2_COPY_740_800 (1-61) x BC062938 (1-5385)
QY 1 ValysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIle 20
Db 2472 GTAAAGTGTATTATACCGAGCGCTGTACCCCTTTGAATCCGAAAGTCAGATGAGATC 2531
QY 21 ThrTlGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr 40
Db 2532 ACCATCCAGCCAGAGATATAGTCATGTGGATGAAAGCCAGACTGGAGGCCAGATGG 2591
QY 41 LeuGlyGlyGluLeuLeuGlyGlySerThrGlyTyrPheProAlaAsnTyrAlaGlyIle 60
Db 2592 CTTGGAGGAGAGCTGAAGGAGAGACGGAGTGGTTCCTGCAACTATGACGAAAGAAATT 2651
QY 61 Pro 61
Db 2652 CCA 2654
RESULT 10
CAS11580 608 bp mRNA linear EST 15-NOV-2002
LOCUS UI-R-FJO-cpy-m-13-0-UI.r1 UI-R-FJO Rattus norvegicus cDNA clone
DEFINITION UI-R-FJO-cpy-m-13-0-UI 5', mRNA sequence.
ACCESSION CAS11580
VERSION CAS11580.1 GI:25002534
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 608)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA library preparation: Dr. M. Bento Soares, University of Iowa
CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 REVERSE.
FEATURES
source location/Qualifiers
1..608
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-FJO-cpy-m-13-0-UI"
/tissue_type="embryo"
/dev_stage="embryo"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-R-FJO"
/note="Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I;
UI-R-FJO is a cDNA library containing the following
tissue(s): rat embryo. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested

with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CATCTCTACT. This library
was created for the University of Iowa Program for Rat
Gene Discovery and Mapping (Val Sheffield, Bento Soares
and Tom Casavant)"
ORIGIN
Alignment Scores:
Pred. No.: 5,17e-31 Length: 608
Score: 329.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.80% Indels: 0
DB: 14 Gaps: 0
US-09-720-934-2_COPY_740_800 (1-61) x CAS11580 (1-608)
QY 2 LysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIleThr 21
Db 135 AAAGTGTGTATTATACCGAGCGCTGTACCCCTTTGAGTCCAGAAATCATGACGATCAC 194
QY 22 TleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyrPleu 41
Db 195 ATCCAGCCAGAGATATAGTCATGTGGATGAAAGCCAGACCGAGATGAGCTT 254
QY 42 GlyGlyGlyLeuLeuLeuGlyGlySerThrGlyTyrPheProAlaAsnTyrAlaGlyIlePro 61
Db 255 GGAGGAGAACTGAAGGAGAGACGAGATGGTTCCTGCAACTATGACGAAAGAAATTCCA 314
RESULT 11
CBS27154 703 bp mRNA linear EST 09-JUL-2003
LOCUS UI-M-FVO-cfh-m-16-0-UI.r1 NIH BMAP_FVO Mus musculus cDNA clone
DEFINITION IMAGE:6849689 5', mRNA sequence.
ACCESSION CBS27154
VERSION CBS27154.1 GI:29360627
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 703)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE NIH-MGC
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA library preparation: Dr. M. Bento Soares, University of Iowa
CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
FEATURES
source location/Qualifiers
1..703
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6849689"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_FVO"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGCAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.:	8.33e-30	Length:	703
Score:	320.00	Matches:	59
Percent Similarity:	96.72%	Conservative:	0
Best Local Similarity:	96.72%	Mismatches:	2
Query Match:	96.10%	Indels:	0
DB:	14	Gaps:	0

US-09-720-934-2_COPY_740_800 (1-61) x CB527154 (1-703)

QY 1 VallyValValTYTYTARGAlaLeuTYrProPheGluSerArgSerHisAspGluile 20
514 GTAAAGTGTATATACCGAGCGCTGACCCCTTGATCCAAAGTCAGAGATGC 573
DB 21 ThrllleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTYr 40
574 ACCATCCAGCCAGAGATATATGTCATGTCGATGATGAAGCCAGACTGAGAGACCAGATG 633
QY 41 LeuGlyGlyGluLeuLysGlyLysThrGlyTYrPheProAlaAsnTYrAlaGluLysile 60
634 CTGGAGAGAGCTGANAAGGAGACGAGATGTTCCCTGCANACTATGCGAAGAAAGATT 693
QY 61 Pro 61
694 CCA 696

RESULT 12

CF735070

LOCUS CF735070 673 bp mRNA linear EST 10-OCT-2003
DEFINITION UT-M-HB0-ckc-d-09-0-UT.r1 NIH BMAP_HB0 Mus musculus cDNA clone
IMAGE:30615536 5', mRNA sequence.

ACCESSION CF735070.1 GI:37631406

VERSION CF735070.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 673)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mouse1.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.
Location/Qualifiers
1. 673
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"

FEATURES

source

/clone="IMAGE:30615536"
/tissue_type="whole eye"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_idb="NIH BMAP HB0"
/note="Organ: Eye; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TTTATGAGT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

ORIGIN

Alignment Scores:

Pred. No.:	1.05e-29	Length:	673
Score:	319.00	Matches:	59
Percent Similarity:	96.72%	Conservative:	0
Best Local Similarity:	96.72%	Mismatches:	2
Query Match:	95.80%	Indels:	0
DB:	14	Gaps:	0

US-09-720-934-2_COPY_740_800 (1-61) x CF735070 (1-673)

QY 1 VallyValValTYTYTARGAlaLeuTYrProPheGluSerArgSerHisAspGluile 20
483 GTAAAGTGTATATACCGAGCGCTGACCCCTTGATCCAAAGTCAGAGATGC 542
DB 21 ThrllleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTYr 40
543 ACCATCCAGCCAGAGATATATGTCATGTCGATGATGAAGCCAGACTGAGAGACCAGATG 602
QY 41 LeuGlyGlyGluLeuLysGlyLysThrGlyTYrPheProAlaAsnTYrAlaGluLysile 60
603 CTGGAGAGAGCTGANAAGGAGACGAGATGTTCCCTGCANACTATGCGAAGAAAGATT 662
QY 61 Pro 61
663 CCA 665

RESULT 13

AA492783

LOCUS AA492783 443 bp mRNA linear EST 25-JUN-1997
DEFINITION v177907.r1 StrataGene mouse testis (#937308) Mus musculus cDNA
clone IMAGE:918500 5', mRNA sequence.

ACCESSION AA492783.1 GI:2222345

VERSION AA492783.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 443)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Stepec,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.

TITLE The Washu-HNMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
Washu-HNMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LBNL; contact the
IMAGE Consortium (info@image.lbnl.gov) for further information.
MWI:530516

Seq primer: -28ml3 rev1 ET from Amerisham
High quality sequence stop: 358.

FEATURES

source

Location/Qualifiers

```

1..443
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CD-1"
/db_xref="taxon:10090"
/clone="IMAGE:918300"
/sex="males"
/tissue_type="testis"
/dev_stage="10-12 week old"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse testis (#937308)"
/notes="Organ: testis; Vector: pBlueScript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Average insert size: 1.0 Kb; Uni-ZAP XR Vector;
-5' adaptor sequence: 5' GAATTCGACAGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"

```

ORIGIN

Alignment Scores:

Pred. No.:	Length:	443
Score:	317.00	Matches: 59
Percent Similarity:	98.36%	Conservative: 1
Best Local Similarity:	96.72%	Mismatches: 1
Query Match:	95.20%	Indels: 0
DB:	9	Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AAA92783 (1-443)

```

QY      1 VallysValValYrYrArgAlaLeuYrProPhgIuseRgSerHisAspGluIle 20
        |||
DB      148 GTRAAAGTGTATATATACCGAGCGCTGTACCCCTTTGAATCCGAAAGTCACGATGAGATC 207
        |||
QY      21 ThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr 40
        |||
DB      208 ACCATCCAGCCAGAGATATAGTCATGTGATGAAAGCCAGACTGAGAGCCAGAGATGC 267
        |||
QY      41 LeuGlyGlyGluLeuLeuGlyGlyGlyThrGlyTyrPheProAlaAsnTyrAlaGluValIle 60
        |||
DB      268 GTTGGAGGAGAGCTGAAGGAGAGACGGATGTTCTCCGCAACTATGACGAAAGATTT 327
        |||
QY      61 Pro 61
        |||
DB      328 CCA 330

```

RESULT 14
CF738758 600 bp mRNA linear EST 10-OCT-2003
LOCUS CF738758
DEFINITION UI-M-HDO-cHu-c-22-0-UI-r1 NIH_BMAP_HDO Mus musculus cDNA clone
IMAGE:30610533 5', mRNA sequence.

ACCESSION CF738758
VERSION CF738758.1 GI:37635095
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NIH-Base 1 to 600
NIH-Base http://mc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs@remail.nih.gov

Tissue Procurement: Dr. James Lin University of Iowa

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

COMMENT

JOURNAL

TITLE

AUTHORS

REFERENCE

Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouse1.html>

This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

FEATURES

source

Location/Qualifiers

```

1..600
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30610533"
/tissue_type="whole eye"
/dev_stage="embryo 12.5, 13.5, 14.5 dpc"
/lab_host="DH10B (r1 phage resistant)"
/clone_lib="NIH BMAP HD0"
/notes="Organ: Eye; Vector: pYX-Asc; Site_1: EcoR I;
Site_2: Not I; The library was constructed according
to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is TTAATGAGT. This library was created for the University
of Iowa Brain Anatomy Project (BMAP); 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."
```

ORIGIN

Alignment Scores:

Pred. No.:	Length:	600
Score:	302.00	Matches: 55
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	90.69%	Indels: 0
DB:	14	Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x CF738758 (1-600)

```

QY      7 ArgAlaLeuYrYrProPhgIuseRgSerHisAspGluIleThrIleGlnProGlyAsp 26
        |||
DB      2 CGAGCGCTGTACCCCTTTAATCCAGAGTCAAGATGAGATCCATCCAGCCAGAGATG 61
        |||
QY      27 IleValMetValAspGluSerGlnThrGlyGluProGlyTyrPheGlyGlyGluLeuLeu 46
        |||
DB      62 ATATGATGTGTGATTAAGCCAGACTGAGAGACCCGAGATGCTTGAGAGAGAGCTGAAA 121
        |||
QY      47 GlyGlyThrGlyTyrPheProAlaAsnTyrAlaGluValIlePro 61
        |||
DB      122 GGAAGACGGAGATGTTCCCTGCAACTATGCAAGAAAGATTCCA 166
        |||

```

RESULT 15
AV669700 551 bp mRNA linear EST 22-SEP-2000
LOCUS AV669700 OLHNI cell line cDNA library (Olb) Oryzias latipes cDNA
DEFINITION clone Olb10.07g similar to adaptor protein intersectin (African
clawed frog), mRNA sequence.

ACCESSION AV669700
VERSION AV669700.1 GI:9935498
KEYWORDS EST.
SOURCE Oryzias latipes (Japanese medaka)
ORGANISM Oryzias latipes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

NIH-Base 1 to 551

Naruse, K., Tanaka, M., Shima, A. and Mitani, H.

Medaka EST Project in University of Tokyo

Unpublished (2000)

COMMENT

JOURNAL

TITLE

AUTHORS

REFERENCE

COMMENT

Contact: Kiyoshi Naruse
Department of Biological Sciences
Graduate School of Science, University of Tokyo
Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan
Tel: 81-3-5841-4443
Fax: 81-3-5841-4410
Email: naruse@biol.s.u-tokyo.ac.jp
This clone was isolated from OLNI cell line cDNA library (OLb) 5'
end sequences.

FEATURES

Location/Qualifiers
1..551
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="HNI"
/db_xref="taxon:8090"
/clone="OLb10.07g"
/clone_lib="OLNI cell line cDNA library (OLb)"

ORIGIN

Alignment Scores:

Pred. No.:	1.46e-27	Length:	551
Score:	301.00	Matches:	51
Percent Similarity:	98.36%	Conservative:	9
Best Local Similarity:	83.61%	Mismatches:	1
Query Match:	90.39%	Indels:	0
DB:	9	Gaps:	0

US-09-720-934-2_COPY_740_800 (1-61) * AV669700 (1-551)

QY	1	ValiysValiValYrYrYrArgAlaLeuTYrProPheGluSerArgSerHisAspGluIle	20
Db	161	CTAAAGTGGTGTATACCGGCGCTGTACCCCTTGTATGCCCGAGCCAGATGATC	220
QY	21	ThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr	40
Db	221	AGCATGTACTGTGAGACCTTATCATGGTGTATGCCAGACGGAGAGCCCGCTGG	280
QY	41	LeuGlyGlyGluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIle	60
Db	281	TTGGGCGGGAGCTCAGAGGTCCGACCGGTGTTCCAGCTAATTATGCAGAACGATT	340
QY	61	Pro	61
Db	341	CCG	343

Search completed: July 1, 2004, 19:48:42
Job time : 1843 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 12:16:50 ; Search time 1345.81 Seconds
(without alignments)
1900.151 Million cell updates/sec

Title: US-09-720-934-2_COPY_908_966
Sequence: 1 VEGLOAQLYPMRAKKDNH.....GEVQKGKMPKSYKVLISG 59

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 segs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-O=/cgen2_1/USPTO.spool_p/US09720934/runat_30062004_064539_13430/app.query.fasta_1.1386
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MIMATCH=0.1 -LOOPT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blcsum62 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09720934_@CEN_1_1_7509_@runat_30062004_064539_13430 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEG_TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
GenEmbl.*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_or:*
21: em_ov:*
22: em_cv:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_man:*
37: em_htg_vtl:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	328	100.0	3231	6 BD205037	BD205037 Isolated
2	328	100.0	3241	9 HSU61166	U61166 Human SH3 d
3	328	100.0	4103	5 AF032118	AF032118 Xenopus 1
4	328	100.0	4321	9 HSM86384	EX538175 Homo sapi
5	328	100.0	5195	6 BD205035	BD205035 Isolated
6	328	100.0	5199	6 BD205033	BD205033 Isolated
7	328	100.0	5287	9 AF064243	AF064243 Homo sapi
8	328	100.0	5381	9 AF114488	AF114488 Homo sapi
9	328	100.0	5458	6 BD205034	BD205034 Isolated
10	328	100.0	6439	9 AF114487	AF114487 Homo sapi
11	328	100.0	7247	9 AF064244	AF064244 Homo sapi
12	328	100.0	52814	9 BS000183	BS000183 Pan trogl
13	328	100.0	100000	9 AP000193	AP000193 Homo sapi
14	328	100.0	114929	9 AP000050	AP000050 Homo sapi
15	328	100.0	125242	9 AP000311	AP000311 Homo sapi
16	328	100.0	142974	2 AC129075	AC129075 Felis cat
17	328	100.0	151516	9 AP000117	AP000117 Homo sapi
18	328	100.0	169774	2 AC138787	AC138787 Sus scrof
19	328	100.0	178801	2 AC121075	AC121075 Canis fam
20	328	100.0	184526	2 AC144367	AC144367 Papio anu
21	328	100.0	194157	2 AC144368	AC144368 Papio anu
22	328	100.0	203423	2 AC137544	AC137544 Gallus ga
23	328	100.0	224721	2 AC139632	AC139632 Bos tauru
24	328	100.0	235426	2 AC121074	AC121074 Canis fam
25	328	100.0	340000	9 AP001718	AP001718 Homo sapi
26	325	99.1	76179	2 AL606725	AL606725 Danio rer
27	325	99.1	104334	5 AL606751	AL606751 Zebrafish
28	325	99.1	163197	5 BX005416	BX005416 Zebrafish
29	325	99.1	182740	2 AC138436	AC138436 Danio rer
30	325	99.1	225535	2 BX470235	BX470235 Danio rer
31	323	98.5	3723	10 AF132478	AF132478 Mus muscu
32	323	98.5	3812	10 AF132672	AF132672 Rattus no
33	323	98.5	4025	10 AF127798	AF127798 Rattus no
34	323	98.5	5145	10 AF132481	AF132481 Mus muscu
35	323	98.5	175319	10 AC134837	AC134837 Mus muscu
36	323	98.5	181782	2 AC116970	AC116970 Rattus no
37	323	98.5	217131	10 AC126053	AC126053 Mus muscu
38	323	98.5	278501	2 AC098552	AC098552 Rattus no
39	323	98.5	279148	2 AC123507	AC123507 Rattus no
40	326	78.0	113146	2 AC138440	AC138440 Tetradon
41	256	78.0	127051	2 AC114895	AC114895 Tetradon
42	251	76.5	1391	9 AF001630	AF001630 Homo sapi
43	251	76.5	2004	9 AF038189	AF038189 Homo sapi
44	251	76.5	2873	6 AR175271	AR175271 Sequence
45	251	76.5	4447	6 AX428899	AX428899 Sequence

RESULT 1

ALIGNMENTS

BD205037
 LOCUS BD205037 3231 bp DNA linear PAT 17-JUL-2003
 DEFINITION Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.
 ACCESSION BD205037
 VERSION BD205037.1 GI:33014807
 KEYWORDS JP 2002511267-A/5.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 3231)
 AUTHORS Korenberg, J.R. and Chen, X.N.
 TITLE Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof
 JOURNAL Patent: JP 2002511267-A 5 16-APR-2002;
 CEDARS SINAI HEALTH SYSTEM ET AL
 COMMENT OS Homo sapiens (human)
 PN JP 2002511267-A/5
 PD 16-APR-2002
 PR 16-APR-1998 JP 2000543610
 PI JULIE R KORENBERG, XIAO NING CHEN
 PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,C12Q1/68,
 PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
 CC Isolated SH3 gene relating to myeloproliferative disorders and leukemia
 CC and utilization thereof.
 FH Key Location/Qualifiers
 FT source 1..3231
 FT 1..3231 /organism='Homo sapiens (human)'.
 ORIGIN
 1..3231
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 Alignment Scores:
 Pred. No.: 2.6e-37 Length: 3231
 Score: 328.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_908_966 (1-59) x BD205037 (1-3231)
 QY 1 ValGUGUGyLeuGlnAlaGlnAlaLeuTyTProTTPaTgaAlaLyLyAspAsnHisLeu 20
 Db 1214 GTGAGGGGGCTACAGCTCAAGCCTTATATCCTTGAGAGCCAAAGCAAGACCACTTA 1273
 QY 21 AsnPhaenLyAsnAspValIleThrValLeuGlnGlnInAspMetTTrpPheGly 40
 Db 1274 AATTTTAACAATAATGATGTCATCACCGTCTGGAACGCAAGACATGTGGTGTGA 1333
 QY 41 GluValGlnGlyGlnLySGlyTTrpPheProLySerTyValLyLeuIleSerGly 59
 Db 1334 GAAGTTCAGGTCAGAGGGTTTGTTCCCAAGTCTTACGTGAACATCATTTCAAGG 1390
 RESULT 2
 LOCUS HSU61166 3241 bp mRNA linear PRI 23-JUL-1996
 DEFINITION Human SH3 domain-containing protein SH3p17 mRNA, complete cds.
 ACCESSION U61166
 VERSION U61166.1 GI:1438932
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 3241)
 AUTHORS Sparks,A.B., Hoffman,N.G., McConnell,S.J., Fowlkes,D.M. and Kay,B.K.
 TITLE Cloning of ligand targets: systematic isolation of SH3 domain-containing proteins
 JOURNAL Nat. Biotechnol. 14 (6), 741-744 (1996)
 MEDLINE 98294438
 PUBMED 9630982
 REFERENCE 2 (bases 1 to 3241)
 AUTHORS Pirozzzi,G., McConnell,S.J., Iveses,A. and Fowlkes,D.M.
 TITLE Direct Submission
 JOURNAL Submitted (18-JUN-1996) CYTOGEN Corp., 307 College Road East, Princeton, NJ 08540, USA
 FEATURES
 source
 1..3241
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="bone marrow"
 37..1599
 /codon_start=1
 /product="SH3 domain-containing protein SH3p17"
 /protein_id="AAC50592.1"
 /db_xref="GI:1438933"
 /translation="MEAEKRLKKEQERKIIIEKQKEBQRAQERDKQMEHYQOED EHQRPKHEEKLRRESVKRQDGEKQEOAKGLRFLPHQHPAKPAVQAWST AEKGLPITSQENKVVYVPALVPESRSHETITOPDIDYVDSQSGEGMIGEL KGTGWFPANAEKIPENEVPAVPKVTDTTSAPKPKALAEPTPLAVTSSEPTTP NNMADFSSTWPTSTNEKRETNDWMAAQPSLTVPSAQLEQRSAPTAATGSSPS VLQGEKEVGLQAOALVPMRAKKNHNFNDVITVLEQDMMWFGVQOQKWFPK SYVKLIIGPIRKTSTMSDSSSPASLIRKVASPAKPVSGEETLAQVIASVTANGPEQ ITLAAGQILIRKKNPGSGWEGELOARKGKQIGMPANVYKLLSPGSKITPPEPK STALAAQGVQIMDYTAQNDELAFNKGQILINVINKEDPMMWKGENVNGVGLPSPNV VKLTITMDPSQQ"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.61e-37 Length: 3241
 Score: 328.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_908_966 (1-59) x HSU61166 (1-3241)
 QY 1 ValGUGUGyLeuGlnAlaGlnAlaLeuTyTProTTPaTgaAlaLyLyAspAsnHisLeu 20
 Db 886 GTGAGGGGGCTACAGCTCAAGCCTTATATCCTTGAGAGCCAAAGCAAGACCACTTA 945
 QY 21 AsnPhaenLyAsnAspValIleThrValLeuGlnGlnInAspMetTTrpPheGly 40
 Db 946 AATTTTAACAATAATGATGTCATCACCGTCTGGAACGCAAGACATGTGGTGTGA 1005
 QY 41 GluValGlnGlyGlnLySGlyTTrpPheProLySerTyValLyLeuIleSerGly 59
 Db 1006 GAAGTTCAGGTCAGAGGGTTTGTTCCCAAGTCTTACGTGAACATCATTTCAAGG 1062
 RESULT 3
 LOCUS AF032118 4103 bp mRNA linear VRT 13-NOV-1998
 DEFINITION Xenopus laevis intersectin mRNA, complete cds.
 ACCESSION AF032118
 VERSION AF032118.1 GI:2642624
 KEYWORDS
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus.
 REFERENCE 1 (bases 1 to 4103)
 AUTHORS Yamahishi,M., Hoffman,N.G., Hardison,N.L., McPherson,P.S., Castagnoli,L., Cesarenti,G. and Kay,B.K.

TITLE Intersectin, a novel adaptor protein with two Eps15 homology and five Src homology 3 domains
JOURNAL J. Biol. Chem. 273 (47), 31401-31407 (1998)
MEDLINE 99030416
PUBMED 9813051
2 (bases 1 to 4103)
AUTHORS Hoffman, N.G., Hardison, N.L., Yamabhai, M. and Kay, B.K.
TITLE Direct Submission
REFERENCE Submitted (30-OCT-1997) Pharmacology, University of Wisconsin, 1300 University Ave, Madison, WI 53706-1532, USA
JOURNAL Location/Qualifiers

FEATURES
source
1..4103
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/cell_type="oocyte"
193..4005
/function="involved in endocytosis"
/note="EH domain and SH3 domain containing protein; similar to Intersectin binding proteins Ibp1, encoded by GenBank Accession Number AF057285, and Ibp2, encoded by GenBank Accession Number AF057286, and mouse Rab/Rip, encoded by GenBank Accession Number AF057287"
/codon_start=1
/product="Intersectin"
/protein_id="AAC73068.1"
/db_xref="GI:2642625"

MAQFGTPEGNGINDIWAIVTEERAKHDQGFHGLKPTAGYITGDQA
RNFFLQSGLPQVLAQIWAIDMANNDRMDQLEFSLAMLIKLGQYPLSLPSNM
LKOPVMPAAVAFVFGMSGIVGIPLAAPVPMPSIPVAGMSPLVSVPTVPLPSN
GAPAVIOHPAFVAFVNSATLPRKSSFGFSVAGSOINTLQKAGSPDPAAPLVAVWAPVS
SSRLKRLQFNSODKTMSSGNLTPQKARTILMOSLSLQSLATIMNISDIDQKLTAE
BFLAHLIDVANSQGLPPLIPPEYIPSPFRVRSGLTMSVSVQRLPEEER
EBQNDADKLVTFEDKREBNPERGMELEKROALBOQRKQELAQERABQERK
EERQOERKROODLEKOLEKOREEERKEIEREAKRELPELQEROLEMER
RROELINONREODIVLVKAKKTELEPELAKDKHOLEGLQDRIOLITOREHT
ESTNKRRLRIAEITHLQOOLQLOESQOLIKMIPEROSLIDOLKQVQNSLHSDILTL
KRALTEKIGRQQLRQDLQBEKEETAKLOEIDVENNOLKELELYNKQFOKQODE
TEKIKOKLEKTSLELDKLEEDKRRMLEODKLMORVQOEERLYNQOEBEKREES
VOKCEVAKPELEIDENKRPFPHPPEPGKLGQIIPMNTEKAPLTIQGVKVVYAL
YFDASVHDEITIEPGDILMVDOSQGEBSGLGELKGTGMFPANVARMPESEBPS
TKPBAETTPAKPVIHVASPVAAPAAFTNTSTNNMNADESSWTPTNTDKVESDMDYT
MAOPSLTVPSAGOHORSAFPATVTVGSSPSPIVGGKVEGLQALVYPMARKND
HINFNKNVITVLEQDDMMFGEVQOKMFPFSYKALISGRKSTSIDTSSEBPA
SLKRVSSPAFKPAIIGEYITSMITYSNEQDITFOQGLIVIKDGMWMTVIEK
TGVPSPNYVRPKDSEAGSGKTSLGKKPEIAQVIAASYAATAPEQLTAPGQLIIR
KNPGGMMWEGELQARGKKRQIGWFPANVYALSPGKSTPTPEPPKTLPTPCQVIG
MDYIKONDELAFSGKGVINLVNKEDPMWKGELNGHGLFESNVYKLTDDMDPSQO
FLGVBPAGGIGATGDRPFLIPFRDGPDLPRAPAPPLASVVMIKFRCTARPCED
MNVKYNIT"

misc_feature 253..498
/note="encodes EHA domain"
misc_feature 850..1101
/note="encodes EHB domain"
misc_feature 2386..2571
/note="encodes SH3A domain"
misc_feature 2881..3057
/note="encodes SH3B domain"
misc_feature 3148..3324
/note="encodes SH3C domain"
misc_feature 3364..3558
/note="encodes SH3D domain"
misc_feature 3607..3786
/note="encodes SH3E domain"

ORIGIN
Alignment Scores:
Pred. No.: 3,4e-37 length: 4103
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x AF032118 (1-4103)
QY 1 valgluglyleuglnalaglnalaleutyprctirpArgAlalyslsyAspAenhlsleu 20
Db 2881 GTAGAGGCGCTTCAGCACAGCGCTGTATCTTGGAGAGCAAGAAGCAACATCTT 2940
QY 21 AsnpeAsnlyAspAspVallethrValleuglGlnGlnaspmetTrrpnegly 40
Db 2941 AATTTTACAAATATATATGTTCACAGGCTTCGAAACAGAGATGTGTGTGGG 3000
QY 41 GluValGlnGlyGlnlysglyTrrpneProlySerTyVallylsrleuiesergly 59
Db 3001 GAAGTTCAAGGTCAAAGGCGCTGTGTTCCCAATCTTATGTAAAGCTTATATCCGT 3057

RESULT 4
HSM806384
LOCUS HSM806384 4321 bp mRNA linear PRI 17-JUN-2003
DEFINITION Homo sapiens mRNA; cDNA DKFZp686J17173 (from clone DKFZp686J17173).
ACCESSION BX538175
VERSION BX538175.1 GI:31874620
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 4321)
Lauber, J., Bahr, A., Mewes, H.W., Weill, B., Amid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.
Direct Submission
Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764

COMMENT
Neuberberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp686J17173) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
Location/Qualifiers

FEATURES
source
1..4321
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="21q22.11"
/clone="DKFZp686J17173"
/tissue_type="human uterus endochel primary cell culture"
/clone_lib="686 (synonym: hicc3). Vector psport1_sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
1..4321
/gene="DKFZp686J17173"
<1..1886
/gene="DKFZp686J17173"
/note="Intersectin long isoform, N-terminus truncated,
differentially spliced"
/codon_start=3
/product="hypothetical protein"
/protein_id="CAD96050.1"
/db_xref="GI:31874621"

/translation="PFFFFFLVVVHLKLAPVYIPLPAPHPAATLPKSSFFSSSG
PSQNTKIQKQSFVAVSPVAVAVQSSRLKXKRLQFNSHDKTMSGHLTPQART
IMOSGLPOALASIMNLSIDIDQDKLTAEFLIHLIDVANSQGLPPLVPEYIP
PSRRVRSQSGSVISSTVDQGLEPEVLEDOQLEKKLPVFEKRENERENAL
ELEKRRQALBOQRKEORLAOLERAEOQRKEERBOERKQLELEKOLEKRELE
QREERKATIEREAKRELRLRQRLERMRNRQRLNQRNQEODIVITKAKKTLF
FEBELANDKRDHQBGLQDTRCLRTTQRLQSEISITNKRRLRIAEITHLQOOLQESQOM
LQGLPEKQILNDQLKQVQNSLHRSLVTLKRLAKELARQLHLDQDVEKETS
KLOEIDIPNNKOLEIREHNKQOLQKQKMEARLQKQERKTIIELEKQKEAORRA
ORDKQMLHVOQDEDEHORPKRIHEEKIKRESYKVKDGEKQEAQDKLGRLEHQ
HDEPAKPAVQAWSTAKGRLTISAQENKVVYALVPPESRSHDEITIEPGDIIWV

polyA_signal 4267..4272
/gene="DKFZp686j17173"
polyA_site 4289
/gene="DKFZp686j17173"

ORIGIN

Alignment Scores:

Pred. No.: 3.61e-37 Length: 4321
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x HSM606384 (1-4321)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTTPArgAlaAlaGlySAspAsnHisLeu 20
DB 2289 GTGAGGGGGCTACAGCTCAAGCCCTATATCTTGAGAGGCCAAGCAACACCACTTA 2348

QY 21 AsnHeaenlySAsnAspValIleThrValLeuGluGlnGlnAspMetTTPTriPheGly 40
DB 2349 AATTTTAACAATAATGATGTCATCACCGCTCGAAGCAAGACATGTGTGTTTGA 2408

QY 41 GluValGlnGlyGlnGlyTyrPheProlySerTyrValIleLeuIleSerGly 59
DB 2409 GAAGTTCAAGGTCAGAGGGTTGGTTCCTCCCAAGCTTACGTGAACCTCATTTCAAGG 2465

RESULT 5 BD205035 5195 bp DNA linear PAT 17-JUL-2003
LOCUS BD205035 Isolated SH3 gene relating to myeloproliferative disorders and
DEFINITION leukemia and utilization thereof.

ACCESSION BD205035
VERSION BD205035.1 GI:33014805
KEYWORDS JP 2002511267-A/3.

SOURCE

ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 5195)
AUTHORS Korenberg,J.R. and Chen,X.N.
TITLE Isolated SH3 gene relating to myeloproliferative disorders and
leukemia and utilization thereof

JOURNAL Patent: JP 2002511267-A 3 16-APR-2002;

COMMENT CEDARS SINAI HEALTH SYSTEM ET AL

OS Homo sapiens (human)

PN JP 2002511267-A/3

PD 16-APR-2002

PF 16-APR-1999 JP 2000543610

PR 16-APR-1999 US 60/082007

PI JULIE R KORENBERG,XIAO NING CHEN

PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,

PC C1201/68,

PC G01N33/68//A61K48/00,C12N15/00,C12N5/00

CC Isolated SH3 gene relating to myeloproliferative disorders and

CC CC and utilization thereof.
FH FH Key Location/Qualifiers
FT FT source 1..5195
/organism="Homo sapiens (human)".

FEATURES
source Location/Qualifiers
1..5195
/organism="Homo sapiens (human)".
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 4.44e-37 Length: 5195
Score: 328.00 Matches: 59

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x BD205035 (1-5195)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTTPArgAlaAlaGlySAspAsnHisLeu 20
DB 2360 GTGAGGGGGCTACAGCTCAAGCCCTATATCTTGAGAGGCCAAGCAACACCACTTA 3019

QY 21 AsnHeaenlySAsnAspValIleThrValLeuGluGlnGlnAspMetTTPTriPheGly 40
DB 3020 AATTTTAACAATAATGATGTCATCACCGCTCGAAGCAAGACATGTGTGTTTGA 3079

QY 41 GluValGlnGlyGlnGlyTyrPheProlySerTyrValIleLeuIleSerGly 59
DB 3080 GAAGTTCAAGGTCAGAGGGTTGGTTCCTCCCAAGCTTACGTGAACCTCATTTCAAGG 3136

RESULT 6 BD205033 5199 bp DNA linear PAT 17-JUL-2003
LOCUS BD205033 Isolated SH3 gene relating to myeloproliferative disorders and
DEFINITION leukemia and utilization thereof.

ACCESSION BD205033
VERSION BD205033.1 GI:33014803
KEYWORDS JP 2002511267-A/1.

SOURCE

ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 5199)
AUTHORS Korenberg,J.R. and Chen,X.N.
TITLE Isolated SH3 gene relating to myeloproliferative disorders and
leukemia and utilization thereof

JOURNAL Patent: JP 2002511267-A 1 16-APR-2002;

COMMENT CEDARS SINAI HEALTH SYSTEM ET AL

OS Homo sapiens (human)

PN JP 2002511267-A/1

PD 16-APR-2002

PF 16-APR-1999 JP 2000543610

PR 16-APR-1999 US 60/082007

PI JULIE R KORENBERG,XIAO NING CHEN

PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,

PC C1201/68,

PC G01N33/68//A61K48/00,C12N15/00,C12N5/00

CC Isolated SH3 gene relating to myeloproliferative disorders and

CC CC and utilization thereof.
FH FH Key Location/Qualifiers
FT FT source 1..5199
/organism="Homo sapiens (human)".
Location/Qualifiers
1..5199
/organism="Homo sapiens (human)".
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 4.45e-37 Length: 5199
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x BD205033 (1-5199)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTTPArgAlaAlaGlySAspAsnHisLeu 20
DB 2329 GTGAGGGGGCTACAGCTCAAGCCCTATATCTTGAGAGGCCAAGCAACACCACTTA 2988

```
Qy 21 AasnPhesAnLysAsnAspValIleThrValLeuGlnGlnAspMetTrpTrpPheGly 40
|||||
Db 2989 AATTTTAAACAAATATGATCATACCGTCTGAAACAGACAACTGCGTGTGGA 3048
|||||
Qy 41 GluValGlnGlyGlnLysGlyTrpPheProLysSerTyrValLysLeuIleSergly 59
|||||
Db 3049 GAAGTTCAAGGTACAGAGGTTGTTCCCAAGTCTTACGTGAAACTCATTTTCAGGG 3105
|||||

RESULT 7
AF064243 5287 bp mRNA linear PRI 21-NOV-1998
LOCUS Homo sapiens intersectin short form mRNA, complete cds.
DEFINITION AF064243
ACCESSION AF064243.1 GI:3859852
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 5287)
AUTHORS Guipponi,M., Scott,H.S., Chen,H., Schebesta,A., Rossier,C. and
Antonarakis,S.E.
TITLES Two isoforms of a human intersectin (ITSN) protein are produced by
JOURNAL brain-specific alternative splicing in a stop codon
MEDLINE Genomics 53 (3), 369-376 (1998)
PUBMED 9901974
9799604
2 (bases 1 to 5287)
AUTHORS Guipponi,M., Scott,H.S., Chen,H., Schebesta,A., Rossier,C. and
Antonarakis,S.E.
TITLES Direct Submission
JOURNAL Submitted (05-MAY-1998) Genetics and Microbiology, CMT, 1 rue
Michel-Servet, Geneva 4 CH-1211, Switzerland
FEATURES
source
1..5287
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
/tissue_type="brain"
/dev_stage="fetus"
107..3769
/codon_start=1
/product="Intersectin short form"
/protein_id="AAC78610.1"
/db_xref="GI:3859853"
/translation="MAQPTPEFGSLIDIAITVERBAKHDOOFHSLKPISGPTGDOA
RNFPSGSGPOPVALAOIWLADNMNDGRMDVDFSIAMKLTIKLQGYLPALPPVM
KOQPVASAPFPKMGATSMPTLVAVPVGSIPTVGMSTLVSSVTAVPPLAN
GAPVATQPLPAFAHPATLTPKSSSFSGSPGSLPTKLOKASFPVASPPVAEMAVP
GSSRLKYRLFNASHDKTMSGHLTGPARITLMOSSLPQQLASINWLSIDDDKLTA
BEFTILAMLEIDVMSGQPLPVLPEPIPPSRVRSGSISVISTVDORLPEEPV
LEDDOOLEKTLPTFEDGKRENFERNGLTEKRQALLLEOKRQERLAOLERAEOE
RKRERROBERKOLELEKOLEKORELEKREPEREAKRRELEKRELEKRELEKRE
RNRROELNQRKEQEDIVYLAKKKTLEFELEALNDKRHQLEGLKQIDTRCLTTRQ
ETESTNKEELNIAETLHQQLQSQQLGLLIPKQTLNQLQKQVQNSLHRSVL
TLKRLAEKELARQHLRDLDVEKEKTRSKLOEIDI PNNQLKELEIHKKQLOKOKS
MAERLKQKEQERKITELEKQKEAORAEQDKWLEVEQOEDHORPKLAEKEL
KRESVKPKQKQKEQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ
KVYRYALYPPESRSHDEITQPGDIIVWKGWVESQTEGEGCWAGGLKGTGAPPA
NYAEKIPENVEVAPYKPYTDSATAPAKLALBETAPLAPLAVTSSESTPPNNADST
WPTSTNEKPTENNDAQAQPSLTIVPSAGLQKRAFTPATVATGSSPSVLSQGRKE
GIQAQALPWRARAKNDNHLNFNNNDVITVLEQDDMWFGVQKGMFPKSYKLSIGP
IRKSTMSDSGSPSLKRVASPAKPVSGSEFTIAMTYESSEGDITLFOCGVIL
VTKKDDWMTGVBKAGVFPENYVRLKDSBEGTAGTKSGKKDEIAQVASTAT
GPEQLIARQGLILRKKNPGSGWEGELQARKKQKQKQKQKQKQKQKQKQKQKQKQ
PPKSTALAAVCQUTGMDYTRQNDDELAFNKQIILNVLNKEDPPMKGEVNGVGLF
PSNYKLTITDMSQQ"
```

```
misc_feature /note="encodes EH domain"
2324..2524
/note="encodes SH3 domain"
misc_feature 2843..3019
/note="encodes SH3 domain"
misc_feature 3110..3286
/note="encodes SH3 domain"
misc_feature 3326..3520
/note="encodes SH3 domain"
misc_feature 3569..3748
/note="encodes SH3 domain"
/note="encodes SH3 domain"

ORIGIN
Alignment Scores:
Pred. No.: 4,53e-37 Length: 5287
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: '9 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x AF064243 (1-5287)

Qy 1 ValGlnGlyLeuGlnAlaGlnAlaLeuTyrProTrpArgAlaLysAspAsnHisLeu 20
|||||
Db 2843 GTGAGGGGCTACAGACTGACCTATATCCTTGAGAGCCAAAAGACACCACTTA 2902
|||||

Qy 21 AasnPhesAnLysAsnAspValIleThrValLeuGlnGlnAspMetTrpTrpPheGly 40
|||||
Db 2903 AATTTTAAACAAATATGATCATACCGTCTGAAACAGACAACTGCGTGTGGA 2962
|||||

Qy 41 GluValGlnGlyGlnLysGlyTrpPheProLysSerTyrValLysLeuIleSergly 59
|||||
Db 2963 GAAGTTCAAGGTACAGAGGTTGTTCCCAAGTCTTACGTGAAACTCATTTTCAGGG 3019
|||||

RESULT 8
AF114488
LOCUS AF114488 5381 bp mRNA linear PRI 16-JUL-2002
DEFINITION Homo sapiens intersectin short isoform (ITSN) mRNA, complete cds.
ACCESSION AF114488
VERSION AF114488.1 GI:4808824
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 5381)
AUTHORS Pucharcos,C., Fuentes,J.J., Casas,C., de la Luna,S., Alcantara,S.,
Arbones,M.L., Soriano,E., Estivill,X. and Pritchard,M.
TITLES Alu-splice cloning of human intersectin (ITSN), a putative
JOURNAL multivalent binding protein expressed in proliferating and
MEDLINE differentiating neurons and overexpressed in Down syndrome
PUBMED Bur. J. Hum. Genet. 7 (6), 704-712 (1999)
99415290
10482960
2 (bases 1 to 5381)
AUTHORS Pucharcos,C., Fuentes,J.J., Pritchard,M. and Estivill,X.
TITLES Direct Submission
JOURNAL Submitted (16-DEC-1998) Medical and Molecular Genetics Center,
Cancer Research Institute, L'Hopital de l'lo., Avia.
Castelldefels km. 2,7, Barcelona 08907, Spain
FEATURES
source
1..5381
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
1..5381
/gene="ITSN"
167..3931
/gene="ITSN"
/codon_start=1
misc_feature /note="encodes EH domain"
767..1936
```

/product="intersectin short isoform"
/protein_id="AA029953.1"
/db_xref="GI:4898825"
/translation="MAQFPFGSLDIMAIVTEERAKHDQGFSLKPSIGFTTGOA
RNFPSGGLPQVLAQIMALAMNDGRMDQVFSIAMKLIKTLQYQALPAPVM
KOOPVASSAPAFMGJASMPPLTAVAPVMSIPIVMSPTLVSVPTAPVPLAN
GAPVLOPAPAPAPATLTKSSFSFGSGLNTRLOKQASFDVAVPVAEMAVP
QSRKLYKROLFPHSHDKTLGSLHGPARTIMOSSLPOLALASTWISDIDQCKLTA
EESTLAWHLIDVMSQPLPVLPEETIPPSFRVRSGLSTVISTSVQRLPEEPV
LEBQOOLERKLEVTEDKKRENFERNGLEERKQALIEQKREQERLAQLAEQE
RKRREOREKROQLLEKQLEKQERLEERREERREIRREAKKELERQLEWE
RNRRELNRKQEDIVLKAKKTLLEBENKQLEKQLEKQLEKQLEKQLEKQLE
EISTSKSRELRTAETLHLOOQESQOMGLRIPEKQIDLOLKQVQCHSRDSIV
TVRALFAKRLAORHLPDOLDEYKTRKSLCEIDIPNNOLKELRIHNNQOIOKKS
MEARLKQKQERKILTELKQKEBAQRQERBKQMLEHVOEDDEHORPKLIEBKU
KRRESVKQKGEERKQEAQDKGLRLEHQHEPAPKPVQA PWSTAERKPLTISAQENV
KVVYRALYPFESRSHDEITIOGDIWMVGEWVDESQTEPGMLGELGKGTGMPA
NVAEKI PENEPAPVPKVDSTAPAPKALRETAPLATVSESPSTPNMMAFSEST
WPTSTKEPTDNMDMAAOPSLTVPAGLORSAETPATATGSSPSVIGQOEKVE
GLAQALYPMRAKDDHMANFNKDVITVLEOQMMMGVEQOGKMGPKSVYKLISGP
IRKSTMSDSSSPASIKRVASPAKPVYSGEFTAMTYTESSEGDILFOQGDVIL
VTKKDGWMTGVDGAGVPSNVYRLKDSGSGTAGKTSLGKKEPIAQLAVIASYLT
GPKQLTLPQQLILIRKQNGWMEGLQARGKROIGMPANVYKLISGTSKIPPT
EPKSTALAAVCQVIEMVDTYAQNDELAFNKGQIIVLNKEDPDMWKGENVQVGLF
PSNVYKLTMDPSQ"

PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10,
PC C12Q1/68,
PC G01N33/68//A61K48/00, C12N15/00, C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and
leukemia
CC and utilization thereof.
FH key Location/Qualifiers
FT source 1..5458
FT Location/Qualifiers
1..5458
/organism="Homo sapiens (human)"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 4.62e-37 Length: 5381
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) X BD205034 (1-5458)
Alignment Scores:
Pred. No.: 4.7e-37 Length: 5458
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

ORIGIN

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyProTPrArgAlaIlyslYsAspAsnHisLeu 20
DB 3003 GTGAGGGGGGTACAGGTCAAGCCCTATATCTTGGAGAGGCCAAAAGACAAACCACTTA 3062

QY 21 AsnPhesnIysAsnAspValIleThrValLeuGluGlnGlnAspMetTPrPheGly 40

QY 21 AsnPhesnIysAsnAspValIleThrValLeuGluGlnGlnAspMetTPrPheGly 40

DB 3065 AATTTTACAAAATGATGTCATCACCGTCTGGAACAGCAAGCAATGTGGTTTGA 3124

DB 3063 AATTTTACAAAATGATGTCATCACCGTCTGGAACAGCAAGCAATGTGGTTTGA 3122

QY 41 GluValGlnGlyIlyGlyTPrPheProIysSerTyValIlyLeuIleSerGly 59

QY 41 GluValGlnGlyIlyGlyTPrPheProIysSerTyValIlyLeuIleSerGly 59

DB 3125 GAAGTCAAGGTCAAGAGGTTGGTCCCAAGCTTACGAAATCATTTACGAGG 3181

DB 3123 GAAGTCAAGGTCAAGAGGTTGGTCCCAAGCTTACGAAATCATTTACGAGG 3179

RESULT 9

RESULT 10

BD205034

AF114487

LOCUS

LOCUS

DEFINITION

DEFINITION

VERSION

VERSION

KEYWORDS

KEYWORDS

SOURCE

SOURCE

ORGANISM

ORGANISM

REFERENCE

REFERENCE

AUTHORS

AUTHORS

TITLE

TITLE

JOURNAL

JOURNAL

COMMENT

FEATURES

OS Homo sapiens (human)

source

PF 16-APR-2002

gene

misc_feature 4649..4819 /note="encodes PH domain"
misc_feature 4895..5143 /note="encodes C2 domain"
ORIGIN

Alignment Scores:
Pred. No.: 6.47e-37 Length: 7247
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x AF064244 (1-7247)

Qy 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTrrPargAlaIlyLysAspAsnHisLeu 20
Db 2843 GTGAGAGGGGCTACAGAGCTCAAGCCCTATATCTTGAGAGCCAAAGAACACCACTTA 2902

Qy 21 AsnPhaenLysAsnAspValIleThrValLeuGluGlnGlnAspMetTrrPheGly 40
Db 2903 AATTTTAAACAAAATGATGTCATCACCGTCTCGAACAGACATGTGTGTGTTGGA 2962

Qy 41 GluValGlnGlyGlnLysGlyTrrPheProLysSerTyrValLysLeuIleSerGly 59
Db 2963 GAAGTTCAAGGTCAAGAGGGTTGGTCCCAAGTCTTACGTGAACATCATTTCAAGG 3019

RESULT 12
BS000183 52814 bp DNA linear PRI 07-OCT-2003
LOCUS
DEFINITION Pan troglodytes chromosome 22 clone:PTB-086H16, map 22, partial
sequence, complete sequence.
ACCESSION BS000183 GI:37537450
VERSION
KEYWORDS HTG.
SOURCE
ORGANISM Pan troglodytes (chimpanzee)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

REFERENCE
AUTHORS 1
TITLE The Chimpanzee Chromosome 22 Sequencing Consortium.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 52814)
AUTHORS Wang, S., Cai, Z., Wang, B., Zheng, H., Zhang, Y., Zhang, X., Zhu, G.,
Lu, G., Fu, G. and Chen, Z.
TITLE Direct Submission
JOURNAL Submitted (26-MAY-2003) Shengyue Wang, Chinese National Human
Genome Center at Shanghai, Genomic Sequencing; No.250 Bibo Road,
Zhang Jiang Hi-Tech Park, Shanghai 201203, CHINA
(E-mail:wangsy@chgc.sh.cn, URL:http://www.chgc.sh.cn,
Tel:86-21-50801919, Fax:86-21-50801922)
The Chinese National Human Genome Center at Shanghai, Shanghai, China;
*Chimpanzee Chromosome 22 Sequencing Consortium consists of:
*Chimpanzee National Human Genome Center at Shanghai, Shanghai, China;
*GFP, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
of Molecular Biotechnology, Jena, Germany; *KerlB Genome Research
Center, Daejeon, Korea;
*Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
*National Institute of Genetics, Mishima, Japan;
*National Yang Ming University Genome Research Center, Taipei,
Taiwan;
*RIKEN Genomic Sciences Center, Yokohama, Japan.

COMMENT
----- Genome Center
Center: Chinese National Human Genome Center at Shanghai Center
code: CHGC
Web site: http://chgc.sh.cn
Contact: wangsy@chgc.sh.cn
----- Project Information
Center project name: The Chimpanzee Chromosome 22 Sequencing Project
Center clone name: PTB-086H16
----- Summary Statistics

Sequencing vector: pUC18, 100% of reads
Chemistry: Dye-terminator Big Dye and ET; 100% of reads Assembly
program: Phrap; version 0.990329
Consensus quality: 52712 bases at least Q40
Consensus quality: 52814 bases at least Q30
Consensus quality: 52814 bases at least Q20
Quality coverage: 9.2x

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >= 30);
an attempt was made to resolve all sequencing problems, such as
compressions and repeats; all regions were covered by at one
plasmid
subclone or more than one M13 subclone;
and the assembly was confirmed by restriction digest.

Source information:
The PR1 chimpanzee BAC library was prepared from DNA isolated from
cultured cells established from the blood of a single male
chimpanzee.
Clones may be obtained from Asao Fujiyama and co-workers
(http://www.gsc.riken.go.jp).
VECTOR: pKS145
Sequence Quality Assessment:
This entry has been annotated with sequence
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in
10,000 bp.

Neighboring clones: PTB-301D17(left) and RP43-082007(right).

FEATURES
Location/Qualifiers
1..52814
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="22"
/clone="PTB-086H16"
/clone_id="PTB1 Chimpanzee BAC"

ORIGIN

Alignment Scores:
Pred. No.: 6.12e-36 Length: 52814
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x BS000183 (1-52814)

Qy 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTrrPargAlaIlyLysAspAsnHisLeu 20
Db 43289 GTGAGAGGGGCTACAGAGCTCAAGCCCTATATCTTGAGAGCCAAAGAACACCACTTA 43348

Qy 21 AsnPhaenLysAsnAspValIleThrValLeuGluGlnGlnAspMetTrrPheGly 40
Db 43349 AATTTTAAACAAAATGATGTCATCACCGTCTCGAACAGACATGTGTGTGTTGGA 43408

Qy 41 GluValGlnGlyGlnLysGlyTrrPheProLysSerTyrValLysLeuIleSerGly 59
Db 43409 GAAGTTCAAGGTCAAGAGGGTTGGTCCCAAGTCTTACGTGAACATCATTTCAAGG 43465

RESULT 13
AP000193 100000 bp DNA linear PRI 08-JAN-2000
LOCUS
DEFINITION Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region,
clone 076C10-f32E9, segment 20/21, complete sequence.
ACCESSION AP000193 GI:4826584
VERSION
KEYWORDS HTG.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 100000)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
 TITLE Homo sapiens 2,083,744bp genomic DNA of 21q22.1 (REGION: D21S26-AML CLONE RANGE: Q7G10-f32E9)
 JOURNAL Published Only in Database (1999)
 REFERENCE 2 (bases 1 to 100000)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (10-MAY-1999) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923, Fax:81-42-778-9924)
 COMMENT E. coli transposon insertion: The present data does not contain E. coli transposon sequences which integrated in the original/previous sequences. We determined the boundary between the insertion and genomic sequences experimentally, removed the insertion sequences, reconstituted the present data. The sequencing project is supported by Japan Science Technology Corporation (JST) and The Institute of Physical and Chemical Research (RIKEN).
 FEATURES
 source
 1. 100000
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.1"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,26e-35 Length: 100000
 Score: 328.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-09-720-934-2_COPY_908_966 (1-59) x AP000193 (1-100000)
 QY 1 ValGUGUyUeUGlnAlaGlnAlaLeuTYrProTTPARgAlaLysLysAspAsnHisLeu 20
 DB 46054 GTGAGGGGCTCAAGCTCAAGCCCTATATCTTGGAGAGCCAAAAAGACAACACTTA 46113
 QY 21 AenPheAsnLysAspAspValIleThrValLeuGUGInGlnAspMetTrrTPPheGly 40
 DB 46114 AATTTTAACAAAATGATGTCTACCCGCTCGGAACAGACACTGTGTGGTTTGA 46173
 QY 41 GiUValGInGlyGlnLysGlyTrrPheProLysSerTYrValLysLeuIleSergly 59
 DB 46174 GAAAGTTCAAGGTCAGAGGGTGTTCCTCCCAAGTCTTACGTAAACTCATTTCAAGG 46230
 RESULT 14
 LOCUS AP000050 114929 bp DNA linear PRI 20-NOV-1999
 DEFINITION Homo sapiens genomic DNA, chromosome 21q22.1, segment 21/28,
 complete sequence.
 ACCESSION AP000050
 VERSION AP000050.1 GI:3132360
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 114929)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
 TITLE Homo sapiens genomic DNA, chromosome 21q

JOURNAL Published Only in Database (1998)
 REFERENCE 2 (bases 1 to 114929)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (11-MAY-1998) Masahira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory; Kitasato 1-15-1, Sagamihara 228, Japan (E-mail:hattori@ngc.ims.u-tokyo.ac.jp, Tel:0427-78-9732, Fax:0427-78-9561)
 COMMENT This sequence is conducted by Kitasato University JST sequencing laboratory as a JST sequencing team.
 Principal Investigator: Yoshinuki Sakaki Ph.D.
 phone: +81-3-5449-5622, Fax : +81-3-5449-5445,
 sakaki@ngc.ims.u-tokyo.ac.jp
 Sub-leader: Tadayoshi Shiba Ph.D., Masahira Hattori Ph.D. The sequence is submitted by: Human Genome Sequencing in ALIS project of JST
 Japan Science and Technology Corporation (JST)
 5-3, Yonbancho, Chiyoda-ku, Tokyo 102-0028 Japan
 For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www-alis.tokyo.jst.go.jp/HGS/top.html>) or send email to webmaster@www-alis.tokyo.jst.go.jp.
 FEATURES
 source
 1. 114929
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.1"
 /clone="130N6-149G3f_2"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.48e-35 Length: 114929
 Score: 328.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-09-720-934-2_COPY_908_966 (1-59) x AP000050 (1-114929)
 QY 1 ValGUGUyUeUGlnAlaGlnAlaLeuTYrProTTPARgAlaLysLysAspAsnHisLeu 20
 DB 9467 GTGAGGGGCTCAAGCTCAAGCCCTATATCTTGGAGAGCCAAAAAGACAACACTTA 9526
 QY 21 AenPheAsnLysAspAspValIleThrValLeuGUGInGlnAspMetTrrTPPheGly 40
 DB 9527 AATTTTAACAAAATGATGTCTACCCGCTCGGAACAGACACTGTGTGGTTTGA 9586
 QY 41 GiUValGInGlyGlnLysGlyTrrPheProLysSerTYrValLysLeuIleSergly 59
 DB 9587 GAAAGTTCAAGGTCAGAGGGTGTTCCTCCCAAGTCTTACGTAAACTCATTTCAAGG 9643
 RESULT 15
 LOCUS AP000311 125242 bp DNA linear PRI 20-NOV-1999
 DEFINITION Homo sapiens genomic DNA, chromosome 21q22.1, D21S26-AML region,
 clone: 130N6, complete sequence.
 ACCESSION AP000311
 VERSION AP000311.1 GI:4835680
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 125242)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
 TITLE Homo sapiens 125,242bp genomic DNA of 21q22.1
 JOURNAL Published Only in Database (1999)
 REFERENCE 2 (bases 1 to 125242)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

TITLE
JOURNAL

Fujiyama, A., Yada, T., Totoki, Y. and Sakaki, Y.
Direct Submission
Submitted (13-MAY-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
Kitasato Univ., 1-15-1 Kitasato, Sagami-hara, Kanagawa 228-8555,
Japan (E-mail: hattori@gsc.riken.go.jp,
URL: http://hgp.gsc.riken.go.jp/, Tel: 81-42-778-9923,
Fax: 81-42-778-9924)

COMMENT

The sequence is a part of the data (ACCESSION No. AP000174 -
AP000194).

The sequencing project is supported by Japan Science Technology
Corporation (JST) and The Institute of Physical and Chemical
Research (RIKEN).

FEATURES

Location/Qualifiers
1. .125242
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1"
/clone="130N6"

ORIGIN

Alignment Scores:

Pred. No.:	1.63e-35	Length:	125242
Score:	328.00	Matches:	59
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-720-934-2_COPY_908_966 (1-59) * AP000311 (1-125242)

QY	1	ValGlnGlyLeuGlnAlaGlnAlaLeuTyRProTrpArgAlaLysLysAspAsnHisLeu	20
Db	109467	GTGGAGGGGCTACAGCTCAAGCCCTATATCTTGAGAGCCAAAAAGACAACTTCA	109526
QY	21	AsnPhenylsAsnAspValIleThrValLeuGlnGlnAspMetTrpTrpPheGly	40
Db	109527	AATTTTAACAAAAATGATGATCAACCGTCCTGGACAGCAAGCATGTGTGTGGTGA	109586
QY	41	GluValGlnGlyGlnLysGlyTrpPheProLysSerTyRValLysLeuLeuSerGly	59
Db	109587	GAAGTTCAAGGTCAGAGGTTGGTCCCAAGTCTTACGTGAAGCACTATTTCAGGG	109643

Search completed: July 1, 2004, 16:15:13
Job time : 1396.81 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:03:00 ; Search time 41.3138 Seconds
(without alignments)
792.522 Million cell updates/sec

Title: US-09-720-934-2_COPY_908_966
Perfect score: 328
Sequence: 1 VAGLQAQALYPRAKKDNHL.....GEVQCKGKMPKSYKYLISG 59

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool -P/US09720934/funat 30062004 064540 13455/app.query.fasta_1.1386
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPOPT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdt
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720934.@CEN 1 1 284 @funat 30062004 064540 13455 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV TIMEOUT=120 -WARN TIMEOUT=0 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/2/ina/PTUS.COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	251	76.5	2873	4	US-08-630-915A-193
2	129.5	39.5	970	4	US-09-833-81-1294
3	125.5	38.3	2955	4	US-09-620-312D-676
4	125.5	38.3	3039	4	US-09-620-312D-675
5	124.5	38.0	5325	4	US-10-164-595-17
6	124.5	38.0	5334	4	US-10-164-595-21
7	124.5	38.0	5581	4	US-10-164-595-19
8	120.5	36.7	1458	2	US-08-942-423-68
9	120.5	36.7	2003	2	US-08-942-423-1
10	120.5	36.7	2003	4	US-08-630-915A-25
11	119.5	36.4	1968	4	US-09-023-655-1296
12	117.5	35.8	1867	4	US-08-630-915A-13

13	116.5	35.5	747	4	US-08-630-915A-39	Sequence 39, Appl
14	113.5	34.6	1392	1	US-08-475-894-3	Sequence 3, Appl
15	113.5	34.6	1392	1	US-08-484-710-3	Sequence 3, Appl
16	113.5	34.6	1392	2	US-08-484-709-3	Sequence 3, Appl
17	113.5	34.6	1392	2	US-08-474-697-3	Sequence 3, Appl
18	113.5	34.6	1659	1	US-08-475-894-1	Sequence 1, Appl
19	113.5	34.6	1659	1	US-08-484-710-1	Sequence 1, Appl
20	113.5	34.6	1659	2	US-08-484-709-1	Sequence 1, Appl
21	113.5	34.6	1659	2	US-08-474-697-1	Sequence 1, Appl
22	113.5	34.6	1659	3	US-08-671-354-1	Sequence 1, Appl
23	113	34.5	874	4	US-09-023-655-931	Sequence 931, App
24	111.5	34.0	1640	4	US-09-023-655-650	Sequence 650, App
25	111	33.8	2757	1	US-08-306-691B-48	Sequence 48, Appl
26	111	33.8	2757	5	PCT-US93-06251-79	Sequence 79, Appl
27	111	33.8	2793	1	US-07-646-537B-1	Sequence 1, Appl
28	108	32.9	1414	4	US-09-023-655-1356	Sequence 1356, Ap
29	108	32.9	5330	4	US-09-023-905A-1	Sequence 1, Appl
30	107.5	32.8	2074	4	US-08-630-915A-19	Sequence 19, Appl
31	107.5	32.8	2445	4	US-09-566-921-20	Sequence 20, Appl
32	107.5	32.8	4695	4	US-09-620-312D-379	Sequence 379, App
33	107	32.6	3456	4	US-09-023-905A-5	Sequence 5, Appl
34	107	32.6	3870	1	US-08-138-641-1	Sequence 1, Appl
35	107	32.6	3870	1	US-08-138-641-3	Sequence 1, Appl
36	107	32.6	3893	1	US-08-138-641-3	Sequence 1, Appl
37	107	32.6	3893	1	US-08-138-133-3	Sequence 3, Appl
38	107	32.6	4382	4	US-09-023-905A-3	Sequence 3, Appl
39	106.5	32.5	813	4	US-09-023-655-1238	Sequence 1238, Ap
40	106.5	32.5	933	1	US-09-641-840-1	Sequence 1, Appl
41	106.5	32.5	933	2	US-08-612-857-1	Sequence 1, Appl
42	106.5	32.5	984	4	US-09-765-298A-5	Sequence 5, Appl
43	106.5	32.5	1072	1	US-07-906-349A-2	Sequence 2, Appl
44	106.5	32.5	1072	1	US-08-167-035-5	Sequence 5, Appl
45	106.5	32.5	1072	1	US-08-167-035-49	Sequence 49, Appl

ALIGNMENTS

RESULT 1
US-08-630-915A-193
Sequence 193, Application US/08630915A
Patent No. 6309820
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLES, Dana M.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 48,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090

```
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 193:
SEQUENCE CHARACTERISTICS:
LENGTH: 2873 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-08-630-915A-193

Alignment Scores:
Pred. No.: 2,68e-28
Score: 251.00
Percent Similarity: 86.44%
Best Local Similarity: 74.58%
Query Match: 76.52%
DB: 4
Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x US-08-630-915A-193 (1-2873)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTPrpArgAlaLysLysAspAsnHisLeu 20
DB 645 GTAGAAAACCTTAAAGCAGCAGCCCTTGTCTCTGAGCTGCAAGAAAGATTAACCACTTG 704
QY 21 AsnPheAsnLysAsnAspValIleThrValLeuGlnGlnAspMetTrpTrpPheGly 40
DB 705 AACTTCTAATAACATGACATATTAATCTGTGGAGCAGCAAGAAATGTGGTGG 764
QY 41 GluValGlnGlyGlnLysGlyTrpPheProLysSerTyrValLysLeuIleSerGly 59
DB 765 GAGGTGATGAGGAGAGAGATGGTTTCCAAATCTTATGTCAAGATCATTCCTGG 821

RESULT 2
US-09-833-381-1294/c
Sequence 1294, Application US/0983381
Patent No. 6672186
GENERAL INFORMATION:
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
FILE REFERENCE: 5800-119
CURRENT APPLICATION NUMBER: US/09/833,381
CURRENT FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 09/516,448
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 2050
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1294
LENGTH: 970
TYPE: DNA
ORGANISM: Homo sapiens
US-09-833-381-1294

Alignment Scores:
Pred. No.: 2.22e-10
Score: 129.50
Percent Similarity: 74.00%
Best Local Similarity: 42.00%
Query Match: 39.48%
DB: 4
Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-09-833-381-1294 (1-970)
QY 7 GlnAlaLeuTyrProTPrpArgAlaLysLysAspAsnHisLeuAsnPheAsnLysAsp 26
DB 376 AAGGCTTGATTTCTTTCAAGCCAGCAAGATAGATTGAAATTTGAAAGAGGTGAC 317
QY 27 ValIleThrValLeuGlnLysGlnAspMetTrpTrpPheGlyGlnValGlnGlyGln 45
DB 316 ATTGTATTTATACACGAGAAAGAAAGAGAGATGCTGTTGATCTTTCAATGGGAAA 257
QY 46 LysGlyTrpPheProLysSerTyrValLys 55
```

```
DB 256 AAGGCCATTTCCTCGCCGCTTATGTGAG 227

RESULT 3
US-09-620-312D-676
Sequence 676, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Weinman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
FILE REFERENCE: 784C1P2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 676
LENGTH: 2955
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (169)..(2418)
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(2955)
OTHER INFORMATION: n = a, c, g or g
US-09-620-312D-676

Alignment Scores:
Pred. No.: 4.24e-09
Score: 125.50
Percent Similarity: 66.67%
Best Local Similarity: 40.74%
Query Match: 38.26%
DB: 4
Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-09-620-312D-676 (1-2955)
QY 5 GlnAlaGlnAlaLeuTyrProTPrpArgAlaLysLysAspAsnHisLeuAsnPheAsnLys 24
DB 1618 CGAGCCAAAGGCCCTTGAGCTTGAAGCGGACGACGACGACGAGCTGGCTCCGCAAG 1677
QY 25 AsnAspValIleThrValLeuGlnGlnAspMetTrpTrpPheGlyGlnValGln 43
DB 1678 AACGACATCATCAACAATCGTGTCTCAGAGAGACGAGACTGCTGGGGGAGACTCAAC 1737
QY 44 GluGlnLysGlyTrpPheProLysSerTyrValLysLeuIle 57
DB 1738 GCGCTGCGAGGCTGCTTTCAGCCCAAGTCTGTGAAGTCTCG 1779

RESULT 4
US-09-620-312D-675
Sequence 675, Application US/09620312D
```

```
/ Patent No. 6569662
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Keundi, Vinod
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Xue, Aildong J.
/ APPLICANT: Yang, Yonghong
/ APPLICANT: Wang, Jian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yundong
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Wang, Zhwei
/ APPLICANT: John Tillinghast
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: No. 6569662el Nucleic Acids and
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 784CIP2B
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ CURRENT FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pc_fl_genes Version 1.0
/ SEQ ID NO 675
/ LENGTH: 3039
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (169)..(2502)
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(3039)
/ OTHER INFORMATION: n = a,t,c,o,r,g
US-09-620-312D-675

Alignment Scores:
Pred. No.: 4,41e-09 Length: 3039
Score: 125.50 Matches: 22
Percent Similarity: 66.67% Conservative: 14
Best Local Similarity: 40.74% Mismatches: 17
Query Match: 38.26% Indels: 1
DB: 4 Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-09-620-312D-675 (1-3039)

QY 5 GlnAlaGlnAlaLeuTyrProTTPArgAlaLysLysAspAsnHisLeuAsnPheAsnLys 24
:::|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1702 CGAGCCAAAGCCCTGCTGAGCTTTGAGCGGCAAGAGAGAGAGAGAGAGAGAGAGAG 1761
QY 25 AsnAspValIleThrPheValLeuGlnGlnAsp--MetTyrTrpPheGlyGluValGln 43
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1763 AAGGACATCATCATCATCGTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1821
QY 44 GlyAlaLysGlyTrpPheProLysSerTyrValLysLeuIle 57
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1822 GGCCTCGAGAGGCTGTTTCCAGCCAAAGTTCGTGAGAGTCTG 1863

RESULT 5
US-10-164-595-17
/ Sequence 17, Application US/10164595
/ Patent No. 6657054
/ GENERAL INFORMATION:
/ APPLICANT: Origene Technologies, Inc
/ TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
/ FILE REFERENCE: IU 103 R1
```

```
/ CURRENT APPLICATION NUMBER: US/10/164,595
/ CURRENT FILING DATE: 2002-06-10
/ NUMBER OF SEQ ID NOS: 80
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 17
/ LENGTH: 5325
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (264)..(3473)
/ OTHER INFORMATION:
US-10-164-595-17

Alignment Scores:
Pred. No.: 1.36e-08 Length: 5325
Score: 124.50 Matches: 23
Percent Similarity: 66.67% Conservative: 15
Best Local Similarity: 40.35% Mismatches: 18
Query Match: 37.96% Indels: 1
DB: 4 Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-10-164-595-17 (1-5325)

QY 2 GluGlyLeuGlnAlaGlnAlaLeuTyrProTTPArgAlaLysLysAspAsnHisLeuAsn 21
:::|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3294 AAGGGAGAGAGAGCCCAAGCCCTTATATGATTTCCGAGGAGAGATGAAGATGAATTCC 3353
QY 22 PheAsnLysAsnAspValIleThrPheValLeuGlnGln--GlnAspMetTrpTrpPheGly 40
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3354 TTCAGGCTCGAGATATATATACAGAGCTGGAATCTGTAGATCATACATGAGATGGA 3413
QY 41 GluValGlnGlyGlnLysGlyTrpPheProLysSerTyrValLysLeuIle 57
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3414 GAACCTATGAGAAATCTGGAATATTTCCCAAAACTACATACAGTTTCTA 3464

RESULT 6
US-10-164-595-21
/ Sequence 21, Application US/10164595
/ Patent No. 6657054
/ GENERAL INFORMATION:
/ APPLICANT: Origene Technologies, Inc
/ TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
/ FILE REFERENCE: IU 103 R1
/ CURRENT APPLICATION NUMBER: US/10/164,595
/ CURRENT FILING DATE: 2002-06-10
/ NUMBER OF SEQ ID NOS: 80
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 21
/ LENGTH: 5334
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (264)..(3482)
/ OTHER INFORMATION:
US-10-164-595-21

Alignment Scores:
Pred. No.: 1.36e-08 Length: 5334
Score: 124.50 Matches: 23
Percent Similarity: 66.67% Conservative: 15
Best Local Similarity: 40.35% Mismatches: 18
Query Match: 37.96% Indels: 1
DB: 4 Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-10-164-595-21 (1-5334)

QY 2 GluGlyLeuGlnAlaGlnAlaLeuTyrProTTPArgAlaLysLysAspAsnHisLeuAsn 21
:::|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3303 AAGGGAGAGAGAGCCCAAGCCCTTATATGATTTCCGAGGAGAGATGAAGATGAATTCC 3362
QY 22 PheAsnLysAsnAspValIleThrPheValLeuGlnGln--GlnAspMetTrpTrpPheGly 40
```

```

Db      3363  TTCAAGCGTGAAGTATATATACAGAGCGTGAATCTGTAGTATGATGAGTGTGGA 34222
Oy      41  GlnValGlnGlyGlnLysGlyTyrPheProLysSerTyrValLysLeuile 57
         |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      3423  GAACCTTATGGGAAATACTGGAATATTCTCCCAAAACTACATACAGATTCTTA 3473
RESULT 7
US-10-164-595-19
/ Sequence 19, Application US/10164595
/ Patent No. 6657054
/ GENERAL INFORMATION:
/ APPLICANT: Origene Technologies, Inc
/ TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
/ FILE REFERENCE: 1U 103 R1
/ CURRENT APPLICATION NUMBER: US/10/164,595
/ NUMBER OF SEQ ID NOS: 80
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 19
/ LENGTH: 5581
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (264)..(3332)
/ OTHER INFORMATION:
US-10-164-595-19

Alignment Scores:
Pred. No.: 1,45e-08 Length: 5581
Score: 124.50 Matches: 23
Percent Similarity: 66.67% Conservative: 15
Best Local Similarity: 40.35% Mismatches: 18
Query Match: 37.96% Indels: 1
DB: 4 Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-10-164-595-19 (1-5581)
Oy      2  GlnGlyLeuGlnAlaGlnAlaLeuTyrProTyrPArgAlaLysAspAsnHisLeuAsn 21
         |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      3550  AAGGGAGAGAAAGGCCAAAGCCCTATATGATTTCCGAGGGGGAATGAGATGAACTTCC 3609
Oy      22  PheAsnLysAsnAspValIleThrValLeuGlnGln---GlnAspMetTyrTyrPheGly 40
         |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      3610  TTCAAGCGTGAAGATATATATAACAGAGCTGGATCTGTAGATGATGACGTGATGAGTGA 3669
Oy      41  GlnValGlnGlyGlnLysGlyTyrPheProLysSerTyrValLysLeuile 57
         |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      3670  GAACCTTATGGGAAATACTGGAATATTCTCCCAAAACTACATACAGATTCTTA 3720
RESULT 8
US-08-942-423-68
/ Sequence 68, Application US/08942423
/ Patent No. 5891673
/ GENERAL INFORMATION:
/ APPLICANT: Hashimoto, Yasuhiro
/ APPLICANT: Takemoto, Yoshinori
/ TITLE OF INVENTION: Lck Binding Protein
/ NUMBER OF SEQUENCES: 68
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Syntex (U.S.A.) Inc.
/ STREET: 3401 Hillview Ave.
/ CITY: Palo Alto
/ STATE: California
/ COUNTRY: U.S.A.
/ ZIP: 94303
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:

```

```

/ APPLICATION NUMBER: US/08/942,423
/ FILING DATE: 01-OCT-1997
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/362,715
/ FILING DATE: 23-DEC-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Peries, Rohan
/ REGISTRATION NUMBER: 35,752
/ REFERENCE/DOCKET NUMBER: 28260
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 852-1698
/ TELEFAX: (415) 496-3529
/ INFORMATION FOR SEQ ID NO: 68:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1458 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: CDNA
/
/ US-08-942-423-68
/
Alignment Scores:
Pred. No.: 9 28e-09 Length: 1458
Score: 120.50 Matches: 23
Percent Similarity: 66.07% Conservative: 14
Best Local Similarity: 41.07% Mismatches: 18
Query Match: 36.74% Indels: 1
Gaps: 1
2
/
US-09-720-934-2_COPY_908_966 (1-59) x US-08-942-423-68 (1-1458)
/
Qy 3 G1yLeuglnAlaGlnAlaLeuTyrProTPrArgAlaIysIysAspAsnHisLeuAsnPhe 22
Db 1291 GGGATCTCTGCTTAAGCCCTGTATGATTAACAGAGAGGGAGAGCGATGACCTTTCCTT 1350
/
Qy 23 AsnIysAsnAspValIleThrValLeuGluGlnGlnAspMet---TPrTPrPheGlyGlu 41
Db 1351 GATCCAGATGACATCATCTACTGACATTAAGATGGTGGATGAAGGCTGTGTGCGGGGCCAA 1410
/
Qy 42 ValGnGlyGlnIysGlyTPrPheProIysSerTyrValIysLeuIle 57
Db 1411 TGCCGTGGCCACTTTGGACTTTTCCTTCGCAAACTATGTCAGACTCTC 1458
/
RESULT 9
/ US-08-942-423-1
/ Sequence 1, Application US/08942423
/ Patent No. 5891673
/ GENERAL INFORMATION:
/ APPLICANT: Hashimoto, Yasuhiro
/ APPLICANT: Takemoto, Yoshihiro
/ TITLE OF INVENTION: Lock Binding Protein
/ NUMBER OF SEQUENCES: 68
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Syntex (U.S.A.) Inc.
/ STREET: 3401 Hillview Ave.
/ CITY: Palo Alto
/ STATE: California
/ COUNTRY: U.S.A.
/ ZIP: 94303
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent'n Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/942,423
/ FILING DATE: 01-OCT-1997
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/362,715
/ FILING DATE: 23-DEC-1994
/ ATTORNEY/AGENT INFORMATION:
/

```


Db	826	ATAGTCACTCTCATCAATPAGGACTGCATCGACGTAGGCTGTGCGAAGAGAGCTGAAC	885
Qy	44	GlyGlnLysGlyTrpPheProLysSerTyValIysLeuile	57
Db	886	GGCAGACGAGGCGGTGTCCCGATTAAGTTCGTGAAGTTACTT	927

Search completed: July 1, 2004, 19:53:51
 Job time : 46.3138 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:09:20 ; Search time 181.007 Seconds
(without alignments)
1572.471 Million cell updates/sec

Title: US-09-720-934-2_COPY_908_966
Perfect score: 328
Sequence: 1 VEGLOAQLYPMRAKKDNHL.....GEVGGKGMFKSYVKLISG 59

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3163042 seqs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame.p2n.model -DEV=xlp
-Q=/cgn2_6/US07.spool_p/US09720934/runat_30062004_064541_13481/app_query.fasta_1.1386
-DB=Published Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MARRIX-biosum62
-TRANS-human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=Pct -THR MAX=100
-THR MN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09720934_@CGN_1.1.1500_@runat_30062004_064541_13481
-NPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPEXT=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELDP=6 -DELEXT=7

Database : Published Applications_NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
------------	-------	-------	-------	--------	----	-------------

1	251	76.5	206	9	US-09-864-761-30400	Sequence 30400, A
2	251	76.5	458	9	US-09-864-761-13834	Sequence 13834, A
3	251	76.5	2873	9	US-09-879-957-193	Sequence 193, App
4	251	76.5	3746	11	US-09-764-875-176	Sequence 176, App
5	251	76.5	4210	9	US-09-764-868-125	Sequence 125, App
6	251	76.5	5828	13	US-10-398-885A-15	Sequence 15, App1
7	134.5	41.0	2067	16	US-10-264-049-887	Sequence 887, App
8	129.5	39.5	970	9	US-09-833-381-1294	Sequence 1294, App
9	129.5	39.5	2539	16	US-10-104-047-1655	Sequence 1655, App
C 10	125.5	38.3	2803	11	US-09-764-868-115	Sequence 115, App
C 11	125.5	38.3	2955	15	US-10-037-270-676	Sequence 324, App
C 12	125.5	38.3	2955	16	US-10-117-722-675	Sequence 676, App
13	125.5	38.3	2955	16	US-10-037-270-676	Sequence 676, App
14	125.5	38.3	3039	15	US-10-037-270-675	Sequence 675, App
15	125.5	38.3	3039	16	US-10-117-722-675	Sequence 675, App
16	124.5	38.0	389	10	US-09-918-995-23920	Sequence 23920, A
17	124.5	38.0	2688	11	US-09-764-875-198	Sequence 198, App
18	123.5	37.7	581	10	US-09-918-995-14807	Sequence 14807, A
19	121.5	37.0	4668	13	US-10-087-192-1205	Sequence 1205, App
20	121.5	37.0	4760	15	US-10-240-965-1105	Sequence 105, App
21	120.5	36.7	2003	9	US-09-879-957-125	Sequence 25, App1
22	119.5	36.4	1416	16	US-10-388-934-99	Sequence 99, App1
23	119.5	36.4	1968	9	US-09-880-107-3711	Sequence 3711, App
24	119.5	36.4	1968	13	US-10-342-887-1141	Sequence 1141, App
25	119.5	36.4	1968	13	US-10-172-118-1141	Sequence 1141, App
26	119.5	36.4	1968	16	US-10-133-937-91	Sequence 91, App1
27	119.5	36.4	1968	16	US-10-159-563-91	Sequence 91, App1
28	119.5	36.4	1968	17	US-10-641-643-1296	Sequence 1296, App
29	119.5	36.4	4666	16	US-10-159-563-299	Sequence 299, App
30	117.5	35.8	1867	9	US-09-879-957-13	Sequence 13, App1
31	116.5	35.5	747	9	US-09-879-957-13	Sequence 13, App1
32	116.5	35.5	3248	15	US-10-171-581-185	Sequence 185, App
33	116.5	35.5	4053	13	US-10-342-887-1882	Sequence 1882, App
34	116.5	35.5	4053	13	US-10-172-118-1882	Sequence 1882, App
35	115.5	35.2	2278	9	US-09-764-868-131	Sequence 131, App
36	115.5	35.2	2278	11	US-09-764-875-494	Sequence 494, App
37	114.5	34.9	2967	15	US-10-207-655-197	Sequence 197, App
38	113.5	34.6	386	13	US-10-085-783A-37196	Sequence 37196, A
39	113.5	34.6	386	16	US-10-242-535A-37196	Sequence 37196, A
40	113.5	34.6	400	13	US-10-085-783A-22639	Sequence 22639, A
41	113.5	34.6	400	16	US-10-242-535A-22639	Sequence 22639, A
42	113.5	34.6	1215	17	US-10-648-593-61	Sequence 61, App1
43	113.5	34.6	3143	14	US-10-144-621-1	Sequence 1, App1
44	113.5	34.6	3348	13	US-10-342-887-187	Sequence 187, App
45	113.5	34.6	3348	13	US-10-172-118-187	Sequence 187, App

ALIGNMENTS

RESULT 1
US-09-864-761-30400
Sequence 30400, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Weisheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aeonica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359

```

; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30400
; LENGTH: 206
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008073.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: NT HIT: AF182198.1, EVALUE 1.00e-112
; OTHER INFORMATION: SWISSPROT HIT: Q15811, EVALUE 3.00e-23
; OTHER INFORMATION: EST_HUMAN HIT: AW505025.1, EVALUE 1.00e-112
US-09-864-761-30400

Alignment Scores:
Pred. No.: 2,46e-28 Length: 206
Score: 251.00 Matches: 44
Percent Similarity: 86.44% Conservative: 7
Best Local Similarity: 74.58% Mismatches: 8
Query Match: 76.52% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x US-09-864-761-30400 (1-206)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyPProTfPaGAlaLysAspAsnHisLeu 20
Db 10 GTAGAAAACCTTAAGACACAGGCCCTTGTCTCTGACTGCAAAAGATACCACTTG 69
QY 21 AsnPhaenLysAsnAspValIleThrValLeuGlnGlnGlnAspMetTrpTirPheGly 40
Db 70 AACTTCTCAAAACATGACATTTACTGTCTTGAGCAGCAAGAAATGCTGTTGGG 129
QY 41 GluValGlnGlyGlnLysGlyTirPheProLysSerTyTValLysLeuLleSerGly 59
Db 130 GAGGTGATGAGAGAGAGAGATGGTTCCCAATCTTAATGCAATCATTTCTGGG 186

RESULT 2
US-09-864-761-13834
; Sequence 13834, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
```

```

; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13834
; LENGTH: 458
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008073.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
US-09-864-761-13834

Alignment Scores:
Pred. No.: 6,86e-28 Length: 458
Score: 251.00 Matches: 44
Percent Similarity: 86.44% Conservative: 7
Best Local Similarity: 74.58% Mismatches: 8
Query Match: 76.52% Indels: 0
Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x US-09-864-761-13834 (1-458)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyPProTfPaGAlaLysAspAsnHisLeu 20
Db 247 GTAGAAAACCTTAAGACACAGGCCCTTGTCTCTGACTGCAAAAGATACCACTTG 306
QY 21 AsnPhaenLysAsnAspValIleThrValLeuGlnGlnGlnAspMetTrpTirPheGly 40
Db 247 GTAGAAAACCTTAAGACACAGGCCCTTGTCTCTGACTGCAAAAGATACCACTTG 306
```


US-09-720-934-2_COPY_908_966 (1-59) x US-09-764-868-125 (1-4210)

Qy 1 ValGIuGIyLeuGlnAlaGlnAlaLeuTyProTrpArgAlaLysAspAsnHisLeu 20
Db 1022 GTGAGAAACTTAAAGACAGGCGCTTGTCTCGACTGCAAGAAAGATACCACTTG 1081
Qy 21 AsnPhaenLysAsnAspValIleThrValLeuGluGlnGlnAspMetTyrTrpPheGly 40
Db 1082 AACTTCTCAAAACATGACATTAATTAAGTCTTGAGACAGCAAGAAATGTGTGGTTGGG 1141
Qy 41 GIuValGlnGlyGlnLysGlyTyrPheProLysSerTyrValLysLeuIleSerGly 59
Db 1142 GAGGTGATGAGAGAGAGATGTTTCCCAATCTTATGTCAGATCATTCCTGGG 1198

RESULT 6

US-10-398-885A-15
Sequence 15, Application US/10398885A
Publication No. US20040053282A1
GENERAL INFORMATION:
APPLICANT: Sugita, Yuji
APPLICANT: Hashida, Ryoichi
APPLICANT: Ogawa, Kaoru
APPLICANT: Nagasu, Takeshi
APPLICANT: Obayashi, Masaya
APPLICANT: Saito, Hirohisa
APPLICANT: Takahashi, Eiki
TITLE OF INVENTION: Method of Testing For Allergic Diseases
FILE REFERENCE: SHIMIZU-07907
CURRENT APPLICATION NUMBER: US/10/398, 885A
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: PCT/JP01/08937
PRIOR FILING DATE: 2001-10-11
PRIOR APPLICATION NUMBER: JP 2000-314093
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 15
LENGTH: 5828
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
NAME/KEY: CDS
LOCATION: (7) .. (5052)
OTHER INFORMATION:
US-10-398-885A-15

Alignment Scores:

Pred. No.: 1.79e-26
Score: 251.00
Percent Similarity: 86.44%
Best Local Similarity: 74.58%
Query Match: 76.52%
DB: 13

US-09-720-934-2_COPY_908_966 (1-59) x US-10-398-885A-15 (1-5828)

Length: 5828
Matches: 44
Conservative: 7
Mismatch: 8
Indels: 0
Gaps: 0

Qy 1 ValGIuGIyLeuGlnAlaGlnAlaLeuTyProTrpArgAlaLysAspAsnHisLeu 20
Db 2650 GTAGAAACTTAAAGACAGGCGCTTGTCTCGACTGCAAGAAAGATACCACTTG 2709
Qy 21 AsnPhaenLysAsnAspValIleThrValLeuGluGlnGlnAspMetTyrTrpPheGly 40
Db 2710 AACTTCTCAAAACATGACATTAATTAAGTCTTGAGACAGCAAGAAATGTGTGGTTGGG 2769
Qy 41 GIuValGlnGlyGlnLysGlyTyrPheProLysSerTyrValLysLeuIleSerGly 59
Db 2770 GAGGTGATGAGAGAGATGTTTCCCAATCTTATGTCAGATCATTCCTGGG 2826
RESULT 7
US-10-264-049-887

Sequence 887, Application US/10264049
Publication No. US20040005579A1

GENERAL INFORMATION:
APPLICANT: Blise et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P1133P1
CURRENT APPLICATION NUMBER: US/10/264, 049
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/18569
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: US 60/209,467
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 4360
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 887
LENGTH: 2067
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (5) .. (5)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (2058) .. (2058)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (2063) .. (2063)
OTHER INFORMATION: n equals a,t,g, or c
US-10-264-049-887

Alignment Scores:

Pred. No.: 1.06e-09
Score: 134.50
Percent Similarity: 64.15%
Best Local Similarity: 49.06%
Query Match: 41.01%
DB: 16

US-09-720-934-2_COPY_908_966 (1-59) x US-10-264-049-887 (1-2067)

Qy 5 GlnAlaGlnAlaLeuTyProTrpArgAlaLysAspAsnHisLeuAsnPhaenLys 24
Db 308 CAGGTGATGAGATGATGACATCAACCGCGAGATGACATGAGCTGCGCTTCAACAAG 367
Qy 25 AsnAspValIleThrValLeuGluGlnGlnAspMet---TyrTrpPheGlyGluValGln 43
Db 368 GAGCAGATCATCAACGTCTCTCAACAAGAGAGACCTGACTGTTGAGAAAGATCAAT 427
Qy 44 GIuGlnLysGlyTyrPheProLysSerTyrValLysLeu 56
Db 428 GAGCAAGTGGGAGCTTCTTCCATCCATTAATGTGAAGCTG 466

RESULT 8

US-09-833-381-1294/c
Sequence 1294, Application US/09833381
Patent No. US20020132090A1
GENERAL INFORMATION:
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
FILE REFERENCE: 5800-119
CURRENT APPLICATION NUMBER: US/09/833,381
CURRENT FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 09/516,448
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 2050
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1294
LENGTH: 970
TYPE: DNA
ORGANISM: Homo sapiens
US-09-833-381-1294

Alignment Scores:

Pred. No.: 2,22e-09 Length: 970
Score: 129.50 Matches: 21
Percent Similarity: 74.00% Conservative: 16
Best Local Similarity: 42.00% Mismatches: 12
Query Match: 39.48% Indels: 1
DB: 9 Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-09-833-381-1294 (1-970)

QY 7 GlnAlaLeuTyRProTTPaRgAlaLysLysAspAsnHisLeuAsnPhaAsnLysAspAsn 26

Db 376 AAGGCGCTTGTATCTCTTTCAGCCAGCAGATGATGATGATTTGAAAAGGTCAC 317

QY 27 ValIleThrValIleuGln--GlnGlnAspMetTTPheGlyGluValGlnGlyGln 45

Db 316 ATTGTGATTATACACGAGAAAAGAAAGAGATGCTGTTTGATCTTTGAATGGGAAA 257

QY 46 LysGlyTTPheProLysSerTyRValLys 55

Db 256 AAGGCCATTCTTCTGCCGCTTATGTGAG 227

RESULT 9

US-10-104-047-1655
Sequence 1655, Application US/10104047
Publication No. US20030236392A1
GENERAL INFORMATION:

APPLICANT: HELIX RESEARCH INSTITUTE

TITLE OF INVENTION: NO. US20030236392A1e1 full length cDNA

FILE REFERENCE: H1-A0105

CURRENT APPLICATION NUMBER: US/10/104,047

PRIOR FILING DATE: 2002-03-25

NUMBER OF SEQ ID NOS: 4096

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1655

LENGTH: 2539

TYPE: DNA

ORGANISM: Homo sapiens

US-10-104-047-1655

Alignment Scores:

Pred. No.: 7.63e-09 Length: 2539
Score: 129.50 Matches: 21
Percent Similarity: 74.00% Conservative: 16
Best Local Similarity: 42.00% Mismatches: 12
Query Match: 39.48% Indels: 1
DB: 16 Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-10-104-047-1655 (1-2539)

QY 7 GlnAlaLeuTyRProTTPaRgAlaLysLysAspAsnHisLeuAsnPhaAsnLysAspAsn 26

Db 2257 AAGGCGCTTGTATCTTTCAGCCAGCAGATGATGATGATTTGAAAAGGTCAC 2216

QY 27 ValIleThrValIleuGln--GlnGlnAspMetTTPheGlyGluValGlnGlyGln 45

Db 2317 ATTGTGATTATACACGAGAAAAGAAAGAGATGCTGTTTGATCTTTGAATGGGAAA 2376

QY 46 LysGlyTTPheProLysSerTyRValLys 55

Db 2377 AAGGCCATTCTTCTGCCGCTTATGTGAG 2406

RESULT 10

US-09-764-868-115/c
Sequence 115, Application US/09764868
Patent No. US2002016871A1
GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PT232

CURRENT APPLICATION NUMBER: US/09/764,868

CURRENT FILING DATE: 2001-01-17

Prior application data removed - refer to PALM or file wrapper

NUMBER OF SEQ ID NOS: 1510

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 115

LENGTH: 2803

TYPE: DNA

ORGANISM: Homo sapiens

US-09-764-868-115

Alignment Scores:

Pred. No.: 3.42e-08 Length: 2803
Score: 125.50 Matches: 22
Percent Similarity: 66.67% Conservative: 14
Best Local Similarity: 40.74% Mismatches: 17
Query Match: 38.26% Indels: 1
DB: 9 Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-09-764-868-115 (1-2803)

QY 5 GlnAlaLeuTyRProTTPaRgAlaLysLysAspAsnHisLeuAsnPhaAsnLys 24

Db 1186 CGAGCCAGGCGCTGCTGACCTTTGAGCGCAGCAGACGAGCTGGCTTCGCAAG 1127

QY 25 AsnAspValIleThrValIleuGlnGlnGlnAsp--MetTTPheGlyGluValGln 43

Db 1126 AAGCAGCATCATCAATCATGCTGTCTCAGAGAGACGACGCTGGTGGGAGACTCAAC 1067

QY 44 GlyGlnLysGlyTTPheProLysSerTyRValLys 57

Db 1066 GGCTGCGAGGCTGTGTTTCAGCCAACTGTGTGAAGTCTCG 1025

RESULT 11

US-09-764-875-324/c
Sequence 324, Application US/09764875
Publication No. US20040018969A1
GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: P202

CURRENT APPLICATION NUMBER: US/09/764,875

Prior application data removed - consult PALM or file wrapper

NUMBER OF SEQ ID NOS: 1249

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 324

LENGTH: 2803

TYPE: DNA

ORGANISM: Homo sapiens

US-09-764-875-324

Alignment Scores:

Pred. No.: 3.42e-08 Length: 2803
Score: 125.50 Matches: 22
Percent Similarity: 66.67% Conservative: 14
Best Local Similarity: 40.74% Mismatches: 17
Query Match: 38.26% Indels: 1
DB: 11 Gaps: 1

US-09-720-934-2_COPY_908_966 (1-59) x US-09-764-875-324 (1-2803)

QY 5 GlnAlaLeuTyRProTTPaRgAlaLysLysAspAsnHisLeuAsnPhaAsnLys 24

Db 1186 CGAGCCAGGCGCTGCTGACCTTTGAGCGCAGCAGACGAGCTGGCTTCGCAAG 1127

QY 25 AsnAspValIleThrValIleuGlnGlnGlnAsp--MetTTPheGlyGluValGln 43

Db 1126 AAGCAGCATCATCAATCATGCTGTCTCAGAGAGACGACGCTGGTGGGAGACTCAAC 1067

QY 44 GlyGlnLysGlyTTPheProLysSerTyRValLys 57

Db 1066 GGCTGCGAGGCTGTGTTTCAGCCAACTGTGTGAAGTCTCG 1025


```
APPLICANT: Tillinghast, John
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US20030104529A1 Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/10/037,270
PRIOR FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 675
LENGTH: 3039
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (169)..(2502)
NAME/KEY: misc_feature
LOCATION: (1)..(3039)
OTHER INFORMATION: n = a,t,c or g
US-10-037-270-675
```

```
Alignment Scores:
Pred. No.: 3,79e-08 Length: 3039
Score: 125.50 Matches: 22
Percent Similarity: 66.67% Conservative: 14
Best Local Similarity: 40.74% Mismatches: 17
Query Match: 38.26% Indels: 1
DB: 15 Gaps: 1
```

US-09-720-934-2_COPY_908_966 (1-59) x US-10-037-270-675 (1-3039)

```
QY 5 GlnAlaGlnAlaLeuTyrProTArgAlaLysAspAsnHisLeuAsnPhaEnLys 24
    ::::::::::::::::::::
DB 1702 CGAGCCAGAGCCCTGCTGAGCTTTGAGCGGACAGACGACGAGCTGGGCTTCCGCAAG 1761
    ::::::::::::::::::::

QY 25 AsnAspVal11eThrValLeuGlnGlnAsp---MetTyrTrpPheGlyGluValGln 43
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 1762 AACGACATCATCAATCGTCTCAGAAAGACGACGACTGCTGGGTGGGGAGCTCAAC 1821
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

QY 44 GlyGlnLysGlyTrpPheProLysSerTyrValLysLeu1le 57
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 1822 GGCTGTGAGGCTGTGTTTCCAGCCAGACTTCGTGGAAGTCTCTG 1863
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
```

RESULT 15

```
US-10-117-722-675
; Sequence 675, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1 Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 675
; LENGTH: 3039
; TYPE: DNA
```

```
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (169)..(2502)
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(3039)
OTHER INFORMATION: n = a,t,c or g
US-10-117-722-675
```

```
Alignment Scores:
Pred. No.: 3,79e-08 Length: 3039
Score: 125.50 Matches: 22
Percent Similarity: 66.67% Conservative: 14
Best Local Similarity: 40.74% Mismatches: 17
Query Match: 38.26% Indels: 1
DB: 16 Gaps: 1
```

US-09-720-934-2_COPY_908_966 (1-59) x US-10-117-722-675 (1-3039)

```
QY 5 GlnAlaGlnAlaLeuTyrProTArgAlaLysAspAsnHisLeuAsnPhaEnLys 24
    ::::::::::::::::::::
DB 1702 CGAGCCAGAGCCCTGCTGAGCTTTGAGCGGACAGACGACGAGCTGGGCTTCCGCAAG 1761
    ::::::::::::::::::::

QY 25 AsnAspVal11eThrValLeuGlnGlnAsp---MetTyrTrpPheGlyGluValGln 43
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 1762 AACGACATCATCAATCGTCTCAGAAAGACGACGACTGCTGGGTGGGGAGCTCAAC 1821
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

QY 44 GlyGlnLysGlyTrpPheProLysSerTyrValLysLeu1le 57
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 1822 GGCTGTGAGGCTGTGTTTCCAGCCAGACTTCGTGGAAGTCTCTG 1863
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
```

Search completed: July 1, 2004, 20:16:02
Job time : 188.007 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:01:26 ; Search time 1774.84 Seconds
(without alignments)
992.694 Million cells updates/sec

Title: US-09-720-934-2_COPY_908_966
Perfect score: 328
Sequence: 1 VEGIQAOALYFWRAKKDNHL.....GEVGGQXGKWPXKTVKLISG 59

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+g2n.model -DEV=xlp
-O=/gcn2.1/USPTO.spool.P/US09720934/rnat.30062004.064540.13442/app_query.fasta_1.1386
-DB=BST -QFMT=fastp -SUFFIX=rest -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi -LIST=45
-DOCAI=GN=200 -THR SCORE=Pct -THR MAX=100 -THR MIN=0 -AICN=15 -MODE=LOCAL
-OUTFMT=txt -NORM=ext -HEAPSIZE=500 -MAXLEN=2000000000
-USER=US09720934@CGN 1.1 12421 @rnat.30062004.064540.13442 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:
1: em.esbba:*
2: em.esbba:*
3: em.esbba:*
4: em.esbba:*
5: em.esbba:*
6: em.esbba:*
7: em.esbba:*
8: em.esbba:*
9: gb.esb1:*
10: gb.esb1:*
11: gb.esb1:*
12: gb.esb1:*
13: gb.esb1:*
14: gb.esb1:*
15: em.esbba:*
16: em.esbba:*
17: em.esbba:*
18: em.esbba:*
19: em.esbba:*
20: em.esbba:*
21: em.esbba:*
22: em.esbba:*
23: em.esbba:*
24: em.esbba:*
25: em.esbba:*
26: em.esbba:*
27: em.esbba:*
28: gb.esb1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	328	100.0	527 9	AL120063
2	328	100.0	536 12	B1340403
3	328	100.0	553 10	AM643494
4	328	100.0	556 9	AV590689
5	328	100.0	582 10	BP911573
6	328	100.0	651 13	BU272390
7	324	98.8	800 14	CB756509
8	324	98.8	847 14	CB756047
9	323	98.5	597 14	CF172865
10	323	98.5	606 14	CF174208
11	323	98.5	738 12	BM944544
12	323	98.5	5385 11	BC062938
13	320	97.6	434 14	CA360651
14	310	94.5	690 13	BQ179493
15	310	94.5	694 14	CB435046
16	280	85.4	341 10	AM889319
17	276	84.1	330 14	BE763158
18	253	77.1	338 10	R93190
19	252	76.8	845 13	BU455276
20	251	76.5	500 10	AM505025
21	251	76.5	961 12	BM466259
22	250	76.2	481 14	CA886530
23	250	76.2	482 28	CC178744
24	250	76.2	499 14	CA889924
25	250	76.2	522 12	BU095151
26	250	76.2	708 12	BI155542
27	249	75.9	465 14	CB715459
28	245	74.7	635 14	CA327260
29	244	74.4	475 10	AM761705
30	242	74.8	406 28	CC200213
31	235	71.6	343 14	CB693264
32	234	71.3	587 9	AV595297
33	230	70.1	727 14	CF539609
34	223	68.0	650 10	BB656585
35	212	64.6	397 10	BE702057
36	208	63.4	760 13	EX879964
37	202	61.6	521 9	AI303871
38	202	61.6	653 14	CF136055
39	198	60.4	555 9	AV960622
40	197	60.1	527 14	CF729468
41	187	60.1	600 14	CF738758
42	184	56.1	668 14	CF763755
43	180	54.9	628 13	BQ393010
44	179	54.6	492 10	AM206547
45	150	45.7	641 28	BZ399663

ALIGNMENTS

RESULT 1
LOCUS AL120063 527 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZP761L192_r1 761 (synonym: hamy2) Homo sapiens cDNA clone
ACCESSION AL120063
VERSION AL120063.1 GI:59259662
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 527)

AUTHORS Otterwaelder, B., Obermaier, B., Mewes, H.W., Gaassenhuber, J. and Wiemann, S.
 TITLE EST (Otterwaelder, et al.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: MIPS
 MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de
 sequenced by Medigenomix (Martinistedt/Germany) within the CDNA sequencing consortium of the German Genome Project. No sl sequence available.
 This clone (DKFZp761L192) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
 FEATURES
 source Location/Qualifiers
 1..527
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp761L192"
 /tissue_type="adipodala"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_idb="761" (synonym: hamy2)
 /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI"
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.82e-36 Length: 527
 Score: 328.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-09-720-934-2_COPY_908_966 (1-59) x AU120063 (1-527)
 QY 1 ValGIGIYleuGInaGlnAlaLeuTYRProTTPRArgAlaLysLysAspAsnHisLeu 20
 Db 306 GTGAGAGGGGCTCAAGCTCAAGCCCTTATCTTGGAGAGCCAAAGCAACCACTTA 365
 QY 21 AasnPhaSnlysaSnspAvallIethnValleuGInGInaSpMetTTPTPheGly 40
 Db 366 AATTTTAACAAAATATGATGCATCACCGCTCTGGAAGAGACATGTCGTGTTGA 425
 QY 41 GluValGInGlyGInLysGlyTTPheProLysSerTYRValLysLeuIleSerGly 59
 Db 426 GAAGTTCAGGCTCAGAGAGGTGTGTTCCCAAGCTTACGTGAACCTATTTCAGGG 482
 RESULT 2
 BI340403 536 bp mRNA linear EST 30-JUL-2001
 LOCUS B1340403
 DEFINITION 365659 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION B1340403
 VERSION B1340403.1 GI:15033686
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 REFERENCE 1 (bases 1 to 536)
 Fahrentzug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J., Vallet, J., Wise, T., Rohrer, G.A., Perrea, G., Sultana, R., Quakehuber, J. and Keeler, J.W.
 Porcine gene discovery by normalized cDNA-library sequencing and EST cluster assembly
 Mamm. Genome 13 (8), 475-478 (2002)
 JOURNAL MEDLINE
 PUBMED 12226715
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA
 Tel.: 402 762 4366
 Fax: 402 762 4390
 Email: smlt@meatlab.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.
 PCR primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCACGTCAGCAGC
 Plate: 105 row: A column: 4
 Seq primer: ATTAGGACACATATAG.
 FEATURES
 source Location/Qualifiers
 1..536
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_idb="MARC 2P1G"
 /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI; Library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.96e-36 Length: 536
 Score: 328.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0
 US-09-720-934-2_COPY_908_966 (1-59) x B1340403 (1-536)
 QY 1 ValGIGIYleuGInaGlnAlaLeuTYRProTTPRArgAlaLysLysAspAsnHisLeu 20
 Db 322 GTGAGAGGGGCTCAAGCCCAAGCCCTTATCTTGGAGAGCCAAAGCAACCACTTA 381
 QY 21 AasnPhaSnlysaSnspAvallIethnValleuGInGInaSpMetTTPTPheGly 40
 Db 382 AATTTTAACAAAATATGATGCATCACCGCTCTGGAAGAGACATGTCGTGTTGA 441
 QY 41 GluValGInGlyGInLysGlyTTPheProLysSerTYRValLysLeuIleSerGly 59
 Db 442 GAAGTTCAGGCTCAGAGAGGTGTGTTCCCAAGCTTACGTGAACCTATTTCAGGG 498
 RESULT 3
 AM643494 553 bp mRNA linear EST 26-APR-2001
 LOCUS AM643494
 DEFINITION cm30d04.w1 Blackshear/Soares normalized Xenopus egg library Xenopus laevis cDNA clone PBX0129D04 5', mRNA sequence.
 ACCESSION AM643494
 VERSION AM643494.1 GI:7400840
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoes; Pipidae; Xenopodinae; Xenopus.
 REFERENCE 1 (bases 1 to 553)
 Blackshear, P.J., Lai, W.S., Thorn, J.M., Kemington, E.A., Staffa, N.G., Jr., Moore, D.T., Boulard, G.G., Beckstrom-Sternberg, S.M., Touchman, D.W., Bonaldo, M.F. and Soares, M.B.
 The NIHXS Xenopus maternal EST project: interim analysis of the first 13,879 ESTs from unfertilized eggs
 Gene 267 (1), 71-87 (2001)
 JOURNAL MEDLINE
 PUBMED 21211403
 COMMENT 11311557
 Contact: Perry J. Blackshear
 Office of Clinical Research and Laboratory of Signal Transduction
 National Institute of Environmental Health Sciences
 A2-05 NIEHS, 101 Alexander Drive, Research Triangle Park, NC 27709,

USA
Tel: 919 541-4899
Fax: 919 541-4571
Email: black009@niehs.nih.gov
Clone is available through Research Genetics, Inc., 2130 Memorial Parkway, Huntsville, AL 35901
phone 800-533-4363 ext.cdn, fax 256-536-9016 att:cdna, email cdna@resgen.com
DNA Sequencing and analyses performed by National Institutes of Health Intramural Sequencing Center (NISC).
PCR Primers
FORWARD: TGTAAACGACGGCCAGT
BACKWARD: CAGGAACAGCTATGACC
Plate: 0129 row: D column: 04
Seq primer: T7 primer.
Location/Qualifiers
1..553
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="FBX0129D04"
/sex="female"
/tissue_type="unfertilized egg"
/cell_type="unfertilized egg"
/dev_stage="unfertilized egg"
/lab_host="DH10B"
/clone_11b="Blackshear/Soares normalized Xenopus egg library"
/note="Vector: pT73-Pac; Site 1: EcoRI; Site 2: NotI; PolyA-selected mRNA was prepared from unfertilized Xenopus laevis eggs. The library was constructed in the vector pT73-Pac as described in Bonaldo, M.F., Lennon, G. and Soares, M.B. 'Normalization and subcloning: two approaches to facilitate gene discovery', Genome Research 6:791-806, 1996. The first strand synthesis used a NotI-dT8 primer; double stranded cDNAs were ligated to EcoRI adapters, digested with NotI, and directionally cloned into the NotI and EcoRI-digested pT73-Pac vector. The library contained approximately 7.2 X 10⁵ recombinants, with average insert sizes of 1-1.5 kb."

ORIGIN

Alignment Scores:
Pred. No.: 6,24e-36 Length: 553
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x AW643494 (1-553)

QY 1 ValGtGtGtYleuGlnAlaGlnAlaLeuTYrProTTPArgAlaLysLysAspAenHsLseu 20
DB 174 GTTGAAGGCTTATGAGCACAAGCCTTGTATCTTGAAGACCAAGAACCAATCTT 233
QY 21 AasnPhaSnLysAsnAspValIleHrValLeuGlnGlnAspMetTTPTPheGly 40
DB 234 AATTTTAAACAAATATGATGATATCAGGTTCTTGAACAGAGATATGTGTGTTGGG 293
QY 41 GluValGlnGtGlnLysGtYrTPheProLysSerTYrValLysLeuIleSerGly 59
DB 294 GAAGTTCAGAGTCAAAAGGGCTGTTCCTCCAAATCCATGTAAAGCTTATATCCGGT 350

RESULT 4
LOCUS AVS90689 556 bp mRNA linear EST 27-NOV-2001
DEFINITION AVS90689 Bos taurus brain fetus Bos taurus cDNA clone E1BR014A10
5', mRNA sequence.
ACCESSION AVS90689
VERSION AVS90689.1 GI:9701682
KEYWORDS EST.
SOURCE Bos taurus (cow)

ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
1 (bases 1 to 556)
Takesuga, A., Hirotsune, S., Itoh, R., Itohzono, A., Suzuki, H., Aso, H. and Sugimoto, Y.
Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs
Nucleic Acids Res. 29 (22), E108 (2001)
JOURNAL MEDLINE 21570554
PUBMED 11713328
COMMENT
Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@cocoa.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.
Location/Qualifiers
1..556
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="E1BR014A10"
/tissue_type="brain"
/dev_stage="fetus"
/lab_host="DH10B"
/clone_11b="Bos taurus brain fetus"
/note="Vector: pZLI; Site 1: SalI; Site 2: NotI; Poly A was deleted from a NotI site"

ORIGIN

Alignment Scores:
Pred. No.: 6,28e-36 Length: 556
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x AVS90689 (1-556)

QY 1 ValGtGtGtYleuGlnAlaGlnAlaLeuTYrProTTPArgAlaLysLysAspAenHsLseu 20
DB 55 GTTGAAGGCTTATGAGCACAAGCCTTGTATCTTGAAGACCAAGAACCAATCTT 114
QY 21 AasnPhaSnLysAsnAspValIleHrValLeuGlnGlnAspMetTTPTPheGly 40
DB 115 AATTTTAAACAAACCAACCTCATCACCGTACCTGAGACCAAGACCAAGCAATTT 174
QY 41 GluValGlnGtGlnLysGtYrTPheProLysSerTYrValLysLeuIleSerGly 59
DB 175 GAAGTTCAGAGTCAAAAGGGCTGTTCCTCCAAATCCATGTAAAGCTTATATCCGG 231

RESULT 5
LOCUS BF911573 582 bp mRNA linear EST 18-JAN-2001
DEFINITION MR4-UT0090-091100-002-e07 UT0090 Homo sapiens cDNA, mRNA sequence.
BF911573
VERSION BF911573.1 GI:12303031
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 582)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,

Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE
20202663
10737800
PUBMED

COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/ICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR4&ct2=MR4-UT0090-091100-002-e07&ct3=2000-11-09&ct4=1)
Seq primer: puc 18 forward
High quality sequence start: 40
High quality sequence stop: 555.
Location/Qualifiers

FEATURES
source
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/def_stage="Adult"
/clone_id="UT0090"
/note="Organ: uterus tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
Alignment Scores:
Pred. No.: 6,71e-36 Length: 582
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x BP911573 (1-582)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTTPArgAlaIysLysAspAsnHisLeu 20
Db 334 GTGAGAGGGCTACAGCTCAAGGCTTATCTTGAGAGCCAAAAAGACACACTTA 275

QY 21 AspPheAsnLysAsnAspValIleThrValLeuGluGlnGlnAspMetTPTTPheGly 40
Db 274 AATTTTAACAAAATGATGTCATCACCGTCTCGAACAAGACATGCTGTGTTTGA 215

QY 41 GluValGlnGlyGlnLysGlyTTPheProLysSerTyrValLysLeuIleSerGly 59
Db 214 GAAGTTCAGAGTCAGAAAGGTTGGTCCCAAGCTTCACTGAACTCATTTTCAGG 158

RESULT 6
BU272390 651 bp mRNA linear EST 26-NOV-2002
LOCUS 603509535F1 CSROCHN52 Gallus gallus cDNA clone CHEST439f10 5', mRNA
DEFINITION
Sequence.
ACCESSION BU272390
VERSION BU272390.1 GI:25543327
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.
1 (bases 1 to 651)

REFERENCE
AUTHORS
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Pong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)

JOURNAL MEDLINE
22335534
12445392
PUBMED

COMMENT
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers

FEATURES
source
1..651
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, HiseX"
/db_xref="taxon:9031"
/clone="CHEST439f10"
/def_stage="22"
/lab_host="DH10B"
/clone_id="CSROCHN52"
/note="Organ: limbs; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN
Alignment Scores:
Pred. No.: 7.87e-36 Length: 651
Score: 328.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x BU272390 (1-651)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTTPArgAlaIysLysAspAsnHisLeu 20
Db 373 GTGAGAGGACTTCACAGCAGGCTCTGTATCTTGAGAGCAAAAAAGACACACTTA 432

QY 21 AspPheAsnLysAsnAspValIleThrValLeuGluGlnGlnAspMetTPTTPheGly 40
Db 433 AATTTTAACAAAATGATGTCATCACGTTTGAAGCAGCAAGATATGTGTGTTTGA 492

QY 41 GluValGlnGlyGlnLysGlyTTPheProLysSerTyrValLysLeuIleSerGly 59
Db 493 GAAGTTCAGAGCAAAAGGGGTGTTTCCCAATCATATGTGAAGCTTATTTTCAGG 549

RESULT 7
CB756509/c 800 bp mRNA linear EST 16-MAY-2003
LOCUS AGENCOURT 12983401 NICHD XGC Tadi Xenopus laevis cDNA clone
DEFINITION
IMAGE:6877279 3', mRNA sequence.
ACCESSION CB756509
VERSION CB756509.1 GI:29836016
KEYWORDS
SOURCE Xenopus laevis (African clawed frog)

ORGANISM	Xenopus laevis Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus. 1 (bases 1 to 800)
REFERENCE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
AUTHORS	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE	Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Sequenced with vector primer Tissue Procurement: Drs. Donald Brown and Liqun Cai cDNA Library Preparation: CLONTECH cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov plate: LICM315 row: k column: 06 High quality sequence stop: 589.
FEATURES	Location/Qualifiers 1..800
SOURCE	

ORIGIN

Alignment Scores:	
Pred. No.:	3,89e-35
Score:	324.00
Percent Similarity:	100.00%
Best Local Similarity:	96.61%
Query Match:	98.78%
DB:	14
Length:	800
Matches:	57
Conservative:	2
Mismatches:	0
Indels:	0
Gaps:	0

US-09-720-934-2_COPY_908_966 (1-59) x CB756509 (1-800)

QY 1 VALGIUGLYLEUGIALHAGINALALEUYYRPTQTPARQALALYSYLSASAPAMHLSLEU 20
Db 609 GTAGAAAGCTCTTCAGGACCAAGCTTGATCCGTGGAGAGCAAGAAAGGACCAACCATCTC 550
QY 21 AANPIEAENLYASAPSPVALIETETPVALLLEUGLGNLGNAPSPETETPTTPPEGLY 40
Db 549 AATTITTAACAAAATCATATATATACAGGCTCTTGAAACAGACGAAATATGCTGGTTGGG 450
QY 41 GIUVALINGLYGMLNYSGLYTRPHEPPOLYSEETRYVALYLSLEULLESGLY 59
Db 489 GAAAGTTCAGGTCACAAAGGCTGGTITTCACCAATCCATAGTAAAGCTTATAGCCGGT 433

RESULT	8
CB756047/c	
LOCUS	847 bp mRNA linear EST 16-MAY-2003
DEFINITION	AGNCOURT_12983387 NICHD XCC Tad1 Xenopus laevis cDNA clone
IMAGE	:6877279 3', mRNA sequence.
ACCESSION	CB756047
VERSION	CB756047.1 GI:29835552
KEYWORDS	EST.
SOURCE	Xenopus laevis (African clawed frog)
ORGANISM	Xenopus laevis
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1. (bases 1 to 847)	Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Xenopodinae; Xenopus.			
NCI-CGAP	http://www.ncbi.nlm.nih.gov/ncicgap.			
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),				
Tumor Gene Index				
Unpublished (1997)				
Contact: Robert Strausberg, Ph.D.				
Email: cgapbs-remail.nih.gov				
Sequenced with anchor primer				
Tissue Procurement: Drs. Donald Brown and Liqun Cai				
cDNA Library Preparation: CLONTECH				
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)				
DNA Sequencing by: Agencourt Bioscience Corporation				
Clone distribution: NCI-CGAP clone distribution information can be				
found through the I.M.A.G.E. Consortium/LINTL at:				
http://image.lnl.gov				
Plate: L1CM3115	row: k	column: 06		
High quality sequence start: 2				
High quality sequence stop: 746.				
Location/Qualifiers				
1..847				

ORIGIN

Alignment Scores:	
Pred. No.:	4.22e-35
Score:	324.00
Percent Similarity:	100.00%
Best Local Similarity:	96.61%
Query Match:	98.78%
DB:	14
Length:	847
Matches:	57
Conservative:	2
Mismatches:	0
Indels:	0
Gaps:	0

US-09-720-934-2_COPY_908_966 (1-59) X CB756047 (1-847)

QY I ValGIuGIuGIuLeuGI.nAl.naGhAl.nAl.euTrProrTrpAgaI.aLVsLVsAspAsHl.sIeu 20
Db 538 GTAAAGGTCCTTCAGGCCAAGACCTTGATCCGGAGAGCAAAAGAAAGAACACCATCTC 479
QY 21 AsnPhaenLVsAsnaAspValIleThrValLeuGIuGIuGI.nAspMetTrpTrpPheGIy 40
Db 478 AATTTTAACAAAAATGATATTATACACGGTCTTGAACAGCAGATATGTGGTITGG 419
QY 41 GIuValGIuGIuGI.nLVsGIyTrpPheProLVsSerTrValLVsLeuIleSerGIy 59
Db 418 GAAGTTCAGGTCAAAAGGSCtGGITTCACCAATCTATGTAACCTTAATAGCCGGT 362

RESULT 9	CF172865	597 bp	mRNA	linear	EST 25-JUL-2003
LOCUS	B0915A02-5	N1A Mouse Unfractionated Bsg CDNA Library (long 1)	Mus		
DEFINITION	musculus cDNA clone N1A:B0915A02 IMAGE:30473665 5'		mRNA sequence.		
ACCESSION	CF172865				
VERSION	CF172865.1	GI:33282414			
KEYWORDS	EST.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 597)
Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method
Genome Res. 11 (9), 1553-1558 (2001)
JOURNAL
MEDLINE
PUBMED
21429098
11544199
COMMENT
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: B0935 row: A column: 02
Seq primer: M13 Reverse
High quality sequence stop: 597
POLYA=No.

FEATURES
source
1. .597
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taeST:B0915A02-5"
/db_xref="taxon:10090"
/clone="NIA:B0915A02 IMAGE:30473665"
/dev_stage="Unfertilized Egg"
/lab_host="DH10B"
/clone_lib="NIA Mouse Unfertilized Egg cDNA Library (long
1)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
Site_2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]. Total
RNAs were extracted from a pool of 1488 unfertilized eggs.
Double-stranded cDNAs were synthesized with an Oligo(dT)
primer (Invitrogen):
5'-GGAAGTCTTGAATCGAGCGGCCCTTTT-3',
treated with T4 DNA polymerase, and purified by
ethanol-precipitation. The cDNAs were ligated to
Lone-Linker L1-SalI, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 2.5 kb. The library was
constructed by Yulan Piao."

ORIGIN
Alignment Scores:
Pred. No.: 3.55e-35 Length: 597
Score: 323.00 Matches: 58
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.31% Mismatches: 0
Query Match: 98.48% Indels: 0
Gaps: 0
DB: 14
US-09-720-934-2_COPY_908_966 (1-597) * CFI72865 (1-597)

QY 1 ValGUGUyengUlnAgInAlaLeuTyProCTyPargAlaLySAspAsnHisLeu 20
Db 352 GTGGAGGGGCTACAGGCAAGCCCTGTATCCCTGGAGACCAAAAAAGACAAACACTTA 411
QY 21 AasnPhaAnlySAsnAapValIIEThrValleuGluInGlnAspMetTPTTPPhaGly 40
Db 412 AATTTTAAcAAAGTAGCGTCATCACCGTTCTGGAACAGCAAGACATGTGTGTTGGA 471
QY 41 GluValGInGlyGlnLySgLyTTPPhaProLySertyValIlyLeuIleSergly 59

DB 472 GAACTTCAAGTCCAGAGGGTTGGTCCCAAGTCTTACGTAAACTATTCAGGG 528
|||||
RESULT 10
CF174208 606 bp mRNA linear EST 25-JUL-2003
LOCUS CF174208
DEFINITION B0935F02-5 NIA Mouse Unfertilized Egg cDNA Library (long 1) Mus
musculus cDNA clone NIA:B0935F02 IMAGE:30475645 5', mRNA sequence.
ACCESSION CF174208
VERSION CF174208.1 GI:33283757
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS
TITLE
1 (bases 1 to 606)
Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method
Genome Res. 11 (9), 1553-1558 (2001)
JOURNAL
MEDLINE
PUBMED
21429098
11544199
COMMENT
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: B0935 row: F column: 02
Seq primer: M13 Reverse
High quality sequence stop: 606
POLYA=No.

FEATURES
source
1. .606
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taeST:B0935F02-5"
/db_xref="taxon:10090"
/clone="NIA:B0935F02 IMAGE:30475645"
/dev_stage="Unfertilized Egg"
/lab_host="DH10B"
/clone_lib="NIA Mouse Unfertilized Egg cDNA Library (long
1)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
Site_2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]. Total
RNAs were extracted from a pool of 1488 unfertilized eggs.
Double-stranded cDNAs were synthesized with an Oligo(dT)
primer (Invitrogen):
5'-GGAAGTCTTGAATCGAGCGGCCCTTTT-3',
treated with T4 DNA polymerase, and purified by
ethanol-precipitation. The cDNAs were ligated to
Lone-Linker L1-SalI, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 2.5 kb. The library was
constructed by Yulan Piao."

ORIGIN
Alignment Scores:
Pred. No.: 3.62e-35 Length: 606
Score: 323.00 Matches: 58
Percent Similarity: 100.00% Conservative: 1

Best Local Similarity: 98.31% Mismatches: 0
 Query Match: 98.48% Indels: 0
 DB: 14 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x CF174208 (1-606)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrrProTrrArgAlaLysLysAspAsnHisLeu 20
 |||
 DB 217 GTGGAAAGGGCTACAGAGCGAAGCCCTGTATCCTCGAGAGCCAAAAGCAACCACTTA 276
 |||
 QY 21 AsnPhaAsnLysAsnAspValIleThrValLeuGlnGlnAspMetrrPrrPhegly 40
 |||
 DB 277 AATTTTACAAAGTACGATCAGCCGTTCTGGAACAGCAACATGTCGTTTGA 336
 |||
 QY 41 GluValGlnGlyGlnLysGlyTyrrPheProLysSerTyrrValLysLeuIleSergly 59
 |||
 DB 337 GAAGTTCAAGGTGAGAGGGTTGTTCCCAAGTCTTACGTGAACATCAATTTCAGGG 393
 |||

RESULT 11
 BM944544 738 bp mRNA linear EST 14-MAR-2002
 LOCUS UI-M-EH0P-bvr-j-08-0-UI.r1 NIH_BMAP_EH0P Mus musculus cDNA clone
 DEFINITION IMAGE:565975 5', mRNA sequence.

ACCESSION BM944544
 VERSION BM944544
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE NIH-MGC http://mgc.nci.nih.gov/
 AUTHORS NIH-MGC
 TITLE 1 (bases 1 to 738)
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
 http://image.liml.gov

This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: pYX-5.

FEATURES
 source Location/Qualifiers

1..738
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:565975"
 /tissue_type="whole brain"
 /dev_stage="embryo 18.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_EH0P"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA
 tail, is CAGCCAGCAGC. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institute of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 4.8e-35 Length: 738
 Score: 323.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.31% Mismatches: 0
 Query Match: 98.48% Indels: 0
 DB: 12 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x BM944544 (1-738)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrrProTrrArgAlaLysLysAspAsnHisLeu 20
 |||
 DB 407 GTGGAAAGGGCTACAGAGCGAAGCCCTGTATCCTCGAGAGCCAAAAGCAACCACTTA 466
 |||
 QY 21 AsnPhaAsnLysAsnAspValIleThrValLeuGlnGlnAspMetrrPrrPhegly 40
 |||
 DB 467 AATTTTACAAAGTACGATCAGCCGTTCTGGAACAGCAACATGTCGTTTGA 526
 |||
 QY 41 GluValGlnGlyGlnLysGlyTyrrPheProLysSerTyrrValLysLeuIleSergly 59
 |||
 DB 527 GAAGTTCAAGGTGAGAGGGTTGTTCCCAAGTCTTACGTGAACATCAATTTCAGGG 583
 |||

RESULT 12
 BC062938 5385 bp mRNA linear HTC 11-DEC-2003
 LOCUS Mus musculus intersectin (SH3 domain protein 1A), mRNA (cDNA clone
 IMAGE:6839463), containing frame-shift errors.

ACCESSION BC062938
 VERSION BC062938
 KEYWORDS HTC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 5385)

AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Dege,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,B., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
 Mckernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Wollay,K.C., Hale,S., Garcia,A.M., Gay,L.J., Huiyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,
 Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,D.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butlerfield,Y.S., Krzyzanski,M.I., Skalska,U., Smalys,D.E.,
 Scherch,A., Schein,D.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL PUBLISHED 12477932
 REFERENCE 2 (bases 1 to 5385)
 AUTHORS Strausberg,R.
 TITLE Direct Submission
 JOURNAL Submitted (26-NOV-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
 COMMENT Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
 Thomas L. Casavant.

Web site: <http://genome.uiowa.edu>
 Contact: bento-soares@uiowa.edu, tom-casaavant@uiowa.edu
 Bonafide, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A.,
 Fishler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K.,
 Schneetz, T., Smith, C., Snider, E., Tack, D., Trout, K., Walters, J.,
 Casaavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: Place: Row: Column: 0
 This clone has the following problem: frame shifted.

FEATURES

source

Location/Qualifiers
 1..5385
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6839463"
 /tissue_type="Brain"
 /clone_lib="NIH BMAP_GH0"
 /lab_host="DH10B"
 /note="Vector: pYX-ASC"

ORIGIN

Alignment Scores:

Pred. No.: 8.17e-34 Length: 5385
 Score: 323.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.31% Mismatches: 0
 Query Match: 98.48% Indels: 0
 DB: 11 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x BC062938 (1-5385)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyProTTPArgAlaLysLysASPsnHisLeu 20

Db 2976 GTGAGAGGGCTACAGCGCAGACCCCTGTATCCCTGAGAGCCAAAGAACACCACTTA 3035

QY 21 AsnPhaSnlyAsnAspValIleThrValLeuGlnGlnAspMetTTPPhely 40

Db 3036 AATTTTAACTTAAAGTACGATACCCCTTGTGAACAGCAACATGTGTGTGGA 3095

QY 41 GluValGlnGlyGlnLysGlyTTPPhProLysSerTyValLysLeuLisSerGly 59

Db 3096 GAAGTTCAAGCTCAGAGGGTTGGTCCCAAGCTTACGGAACATCATTTTCAGGG 3152

RESULT 13

CA360651

LOCUS

634174 NCCMWA 1RT Oncorhynchus mykiss cDNA clone 1RT108K03_A_F02

5', mRNA sequence.

ACCESSION

CA360651.1 GI:24663989

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

FEATURES

source

Location/Qualifiers
 1..434
 /organism="Oncorhynchus mykiss"
 /mol_type="mRNA"
 /db_xref="taxon:8022"
 /clone="1RT108K03_A_F02"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="NCCMWA 1RT"
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
 library made from pooled tissue from brain, gill, liver,
 spleen, muscle, and kidney."

ORIGIN

Alignment Scores:

Pred. No.: 5.98e-35 Length: 434
 Score: 320.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 2
 Best Local Similarity: 96.61% Mismatches: 0
 Query Match: 97.56% Indels: 0
 DB: 14 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x CA360651 (1-434)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyProTTPArgAlaLysLysASPsnHisLeu 20

Db 140 GTGAGAGGGTTGACAGCTCAGGCTGTATCCCTGTGAGGCGCCAGAGACACCACTTC 199

QY 21 AsnPhaSnlyAsnAspValIleThrValLeuGlnGlnAspMetTTPPhely 40

Db 200 AACTTAAACAGAGCGACGATGATCAGCTGTGAGCAGCAGCAACATGTGTGTGCGGC 259

QY 41 GluValGlnGlyGlnLysGlyTTPPhProLysSerTyValLysLeuLisSerGly 59

Db 260 GAGGTGAGAGGGCGACGCGGCTGTGTTCCCAAGTCTTACGTAAAGCTCATCTGTGT 316

RESULT 14

BO179493

LOCUS

UT-M-EMO-bww-k-03-0-UT r1 NIH_BMAP_EMO Mus musculus cDNA clone

IMAGE:5704058 5', mRNA sequence.

ACCESSION

BO179493

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

/lab host="DH10B (T1 phage resistant)"
 /clone lib="NH1.BMAP.BW0"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA
 tail, is GTGGTGGAA. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."

/db xref="taxon:9913"
 /issue_type="pooled"
 /lab host="DH10B"
 /clone lib="MARC 6BOV"
 /note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
 library made with RNA pooled from multiple tissues
 including liver, lung, hypothalamus, pituitary, and
 placenta/endometrium."

ORIGIN

ORIGIN

Alignment Scores:
 Pred. No.: 3,01e-33 Length: 690
 Score: 310.00 Matches: 58
 Percent Similarity: 98.33% Conservative: 1
 Best Local Similarity: 96.67% Mismatches: 0
 Query Match: 94.51% Indels: 1
 DB: 13 Gaps: 0

Alignment Scores:
 Pred. No.: 3,03e-33 Length: 694
 Score: 310.00 Matches: 58
 Percent Similarity: 98.33% Conservative: 1
 Best Local Similarity: 96.67% Mismatches: 0
 Query Match: 94.51% Indels: 1
 DB: 14 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x BQ179493 (1-690)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTYrProTTPArgAlaLysLysAspAsnHisLeu 20
 |||
 DB 423 GTGGAAGGGCTACAGGCAAGCCCTGTATCCCTGGAGAGCCAAAGAAAGACAACCACTTTA 482

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTYrProTTPArgAlaLysLysAspAsnHisLeu 20
 |||
 DB 120 GTGGAAGGGCTACAGGCAAGCCCTGTATCCCTGGAGAGCCAAAGAAAGACAACCACTTTA 179

QY 20 uAsnPhaenLysAsnAspValIleThrValLeuGlnGlnGlnAspMetTrrTTPheG1 40
 |||
 DB 483 AAATTTTAACAAAAGCAGCTCATCACCGTTCTGGAACACCAAGACATGTGGTGGTGG 542

QY 21 Asn-PhaenLysAsnAspValIleThrValLeuGlnGlnGlnAspMetTrrTTPheG1 40
 |||
 DB 180 AATGTGAACAAAAGCAGCTCATCACCGTTCTGGAACACCAAGACATGTGGTGGTGG 239

QY 40 yGIuValGlnGlyGlnLysGlyTrrPheProLysSerTYrValLysLeuLleSergly 59
 |||
 DB 543 AGAAGTTCAAGGTCAGAGGGTGTTCCTCCCAAGTCTTACGTGAACCTCATTTCAAGG 600

QY 40 yGIuValGlnGlyGlnLysGlyTrrPheProLysSerTYrValLysLeuLleSergly 59
 |||
 DB 240 AGAAGTTCAAGGTCAGAGGGTGTTCCTCCCAAGTCTTACGTGAACCTCATTTCAAGG 297

RESULT 15

Search completed: July 1, 2004, 19:48:48
 Job time : 1780.84 secs

LOCUS CB435046 694 bp mRNA linear EST 25-MAR-2003
 DEFINITION 611770 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.

LOCUS CB435046 694 bp mRNA linear EST 25-MAR-2003
 DEFINITION 611770 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION CB435046
 VERSION CB435046.1 GI:29215685

ACCESSION CB435046
 VERSION CB435046.1 GI:29215685

KEYWORDS EST.

KEYWORDS EST.

SOURCE Bos taurus (cow)

SOURCE Bos taurus (cow)

REFERENCE Smith, T.P.L., Roberts, A.J., Echterkamp, S.E., Chitko-McKown, C.G.,
 Wray, J.E. and Keeler, U.W.

REFERENCE Smith, T.P.L., Roberts, A.J., Echterkamp, S.E., Chitko-McKown, C.G.,
 Wray, J.E. and Keeler, U.W.

A second set of bovine ESTs from pooled-tissue normalized libraries

A second set of bovine ESTs from pooled-tissue normalized libraries

Unpublished (2003)

Unpublished (2003)

CONTACT: Smith TPL

CONTACT: Smith TPL

USDA, ARS, US Meat Animal Research Center

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Tel: 402 762 4366

Fax: 402 762 4390

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called with phred v0.020425.c and

Single pass sequencing. Bases called with phred v0.020425.c and

trimmed with the aid of the trim_alt option. Vector identified with

trimmed with the aid of the trim_alt option. Vector identified with

cross match v0.990329

cross match v0.990329

plate: FQY8036 row: 0 column: 9

plate: FQY8036 row: 0 column: 9

Seq primer: GATATGAGCTCACTATAGG3.

Seq primer: GATATGAGCTCACTATAGG3.

Location/Qualifiers

Location/Qualifiers

1. .694

1. .694

/organism="Bos taurus"

/organism="Bos taurus"

/mol_type="mRNA"

/mol_type="mRNA"

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 12:16:50 ; Search time 1459.86 Seconds
(without alignments)
1900.151 Million cell updates/sec

Title: US-09-720-934-2_COPY_999_1062
Perfect score: 341
Sequence: 1 GGEIAQVIVSYTATGPEQLT.....RGKKRQIGWFPANYKVLSP 64

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 segs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q/cg2.1/USPTO.spool.p/US09720934/runat.30062004.064539.13430/app.query.fasta_1.1386
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPT=0 -LOOPEXT=0
-INITs-bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=200000000
-USER=US09720934.@CEN.1.1.7509.@runat.30062004.064539.13430 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_rod:*
36: em_htg_mem:*
37: em_htg_vrt:*
38: em_sy:*
39: em_hugo_hum:*
40: em_hugo_mus:*
41: em_hugo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	341	100.0	3231	6 BD205037	BD205037 Isolated
2	341	100.0	3241	9 HSU61166	U61166 Human SH3 d
3	341	100.0	3812	10 AF132672	AF132672 Rattus no
4	341	100.0	4321	9 HSM80584	EX38175 Homo sapi
5	341	100.0	5199	6 BD205033	BD205033 Isolated
6	330	96.8	4025	10 AF127798	AF127798 Rattus no
7	330	96.8	5195	6 BD205035	BD205035 Isolated
8	330	96.8	5381	9 AF114488	AF114488 Homo sapi
9	330	96.8	5458	6 BD205034	BD205034 Isolated
10	330	96.8	6439	9 AF114487	AF114487 Homo sapi
11	328	96.2	1996	9 AF180522	AF180522 Homo sapi
12	327	95.9	5287	9 AF064243	AF064243 Homo sapi
13	327	95.9	7247	9 AF064244	AF064244 Homo sapi
14	325	95.3	1133	10 AF169621	AF169621 Mus muscu
15	325	95.3	3723	10 AF132478	AF132478 Mus muscu
16	325	95.3	5145	10 AF132481	AF132481 Mus muscu
17	319	93.5	4103	5 AF032118	AF032118 Xenopus 1
18	318	93.3	78190	2 AC106760	AC106760 Homo sapi
19	318	93.3	179264	2 AC023112	AC023112 Homo sapi
20	303	88.9	171603	9 AC012629	AC012629 Homo sapi
21	258	75.7	76179	2 AL606725	AL606725 Dario rer
22	258	75.7	104334	5 AL606751	AL606751 Zebrafish
23	258	75.7	163197	5 BX005416	BX005416 Zebrafish
24	258	75.7	182740	2 AC138436	AC138436 Dario rer
25	258	75.7	225535	2 BX470235	BX470235 Dario rer
26	253	74.2	747	6 AR175268	AR175268 Sequence
27	253	74.2	1391	9 AF001630	AF001630 Homo sapi
28	253	74.2	2004	9 AF038189	AF038189 Homo sapi
29	253	74.2	2873	6 AR175271	AR175271 Sequence
30	253	74.2	3594	10 AF132480	AF132480 Mus muscu
31	253	74.2	4053	9 HSU61167	U61167 Human SH3 d
32	253	74.2	4447	6 AX428899	AX428899 Sequence
33	253	74.2	4557	9 AF182199	AF182199 Homo sapi
34	253	74.2	4977	10 AF132479	AF132479 Mus muscu
35	253	74.2	5804	10 AK122480	AK122480 Mus muscu
36	253	74.2	5828	6 BD167848	BD167848 Method fo
37	253	74.2	5828	9 AF248540	AF248540 Homo sapi
38	253	74.2	5938	9 AB033082	AB033082 Homo sapi
39	253	74.2	6091	9 AF182198	AF182198 Homo sapi
40	251.5	73.8	134408	2 AC139628	AC139628 Takifugu
41	244.5	71.1	136937	2 AC139627	AC139627 Takifugu
42	242.5	71.1	113146	2 AC138440	AC138440 Tetradon
43	242.5	71.1	127051	2 AC114895	AC114895 Tetradon
44	237	69.5	2408	3 AY069517	AY069517 Drosophi1
45	237	69.5	3750	3 AF054612	AF054612 Drosophi1

RESULT 1

ALIGNMENTS

BD205037	LOCUS	BD205037	3231 bp	DNA	linear	PAT 17-JUL-2003
	DEFINITION	Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.				
	ACCESSION	BD205037				
	VERSION	BD205037.1				
	KEYWORDS	JP 2002511267-A/5.				
	SOURCE	JP 2002511267-A/5.				
	ORGANISM	Homo sapiens (human)				
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
	AUTHORS	1 (bases 1 to 3231) Korenberg,J.R. and Chen,X.N.				
	TITLE	Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof				
	JOURNAL	Patent: JP 2002511267-A 5 16-APR-2002;				
	COMMENT	CENDARS SINAI HEALTH SYSTEM ET AL				
		OS Homo sapiens (human)				
		PN JP 2002511267-A/5				
		PD 16-APR-2002				
		PF 16-APR-1999 JP 2000543610				
		PR 16-APR-1998 US 60/082007				
		PI JULIE R KORENBERG; XIAO NING CHEN				
		PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,				
		PC C12Q1/68,				
		PC G01N33/68//A61K48/00,C12N15/00,C12N5/00				
		CC Isolated SH3 gene relating to myeloproliferative disorders and				
		leukemia				
		and utilization thereof.				
		Location/Qualifiers				
		1..3231				
		/organism="Homo sapiens"				
		/mol_type="genomic DNA"				
		/db_xref="taxon:9606"				
	ORIGIN					
	Alignment Scores:					
	Pred. No.:	6.94e-33				
	Score:	341.00				
	Percent Similarity:	100.00%				
	Best Local Similarity:	100.00%				
	Query Match:	100.00%				
		6				
		Gaps:				
		0				
	US-09-720-934-2_COPY_999_1062 (1-64)	x BD205037 (1-3231)				
	QY	1 G1yG1ug1u1lea1aG1nval1lea1aer1yrr1ha1a1rrg1yrr1rog1uG1nleuthr				
	DB	1487 GGAAGAAGAAATGGCCAGGTAAATTCCTCAATACCCGCCACCGCCCGACGACCACTACT				
	QY	21 leua1aer1rog1yG1nleu1le1p1le1ar1g1y1s1a1n1p1rog1yG1y1rr1rr1r1g1uG1y				
	DB	1547 CTCGCCCTGGTGTACGTGATTTTATCGCAAAAAAGAACCCAGGTGATG1G1GGAGGA				
	QY	41 G1u1e1uG1n1a1a1rrg1y1y1s1y1a1rG1n1le1g1y1rr1p1e1p1ro1a1aer1yrr1val1ys				
	DB	1607 GAGCTGAGAGCACGTGGGAAAAACCGCCAGATAGGCTGGTTCACGCTAATTAATGTAAG				
	QY	61 leu1e1u1e1r1p1ro				
	DB	1667 CTTCTTAAGCCCT 1678				
	RESULT 2					
	LOCUS	HSU61166	3241 bp	mrna	linear	PRI 23-JUL-1996
	DEFINITION	Human SH3 domain-containing protein SH3P17 mRNA, complete cds.				
	ACCESSION	U61166				
	VERSION	U61166.1				
	KEYWORDS	GI:1438932				

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 3241) Sparks,A.B., Hoffman,N.G., McConnell,S.J., Fowlkes,D.M. and Kay,B.K. Cloning of ligand targets: systematic isolation of SH3 domain-containing proteins Nat. Biotechnol. 14 (6), 741-744 (1996)
JOURNAL MEDLINE	Nat. Biotechnol. 14 (6), 741-744 (1996)
PUBMED	98294438
REFERENCE TITLE	2 (bases 1 to 3241) Pitozzi,G., McConnell,S.J., Uveges,A. and Fowlkes,D.M. Direct Submission Submitted (18-JUN-1996) CTOGEN Corp., 307 College Road East, Princeton, NJ 08540, USA
JOURNAL	Location/Qualifiers
FEATURES	1..3241
source	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/tissue_type="bone marrow"
CDS	37..1599
	/codon_start=1
	/product="SH3 domain-containing protein SH3P17"
	/protein_id="AAC50592.1"
	/db_xref="GI:1438933"
	/translation="MEARRKKKOEKRIELEKOEKRAORAEPRDKWLEHVQED EHQRPRLHEBKIKRESYKKKODGEKKGOEKDRIGRIFPHOEPARVAPMST AEGKPRTISAQENKVIVRYALYPESRSKHDELTIQDILVMDESQTGSPWLGGEL KKGTPMPNPAVAEIKPENEPVAPVPKPTDTSAPAPLALREPAIPALYSBSPSTP NNMADESWPTSTINEKETDNMDAMWAQSLTVPSAGLRORSATFPATATGSBP VLQCGKEVGLQAQALPYMPAKKDNIHPKNVI TVLEQDDMMVEGVGGOGWPFK SVKLISGPRIKSTSMDSSESSESPASIKRVSPAARKVVSAGEIAQVIATATGPQC LTAAPOGLIIRKMTDPCGMWEGLRQKRGROIGMPNPANYVKLLSBCTSKITPEPK STALAACQYTKMTDTTAQNDBELAFKQGINVLNKEDPPMKKEVNGOVGLFPSNY VKLTIDMDFSOQ"
ORIGIN	
Alignment Scores:	
Pred. No.:	6,96e-33 length: 3241
Score:	341.00 Matches: 64
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatch: 0
Query Match:	100.00% Indels: 0
DB:	Gaps: 0
US-09-720-934-2_COPY_999_1062 (1-64) x HSU61166 (1-3241)	
OY	1 GYGUGLUULEAGINVALILEALASERTYTRThAlathrglyProGIUGInLeuthr 20
Db	1159 GGAGAAMAATTGCCAGGTATTATTCCTCAWACCCGCCACCGCCCCCGAGACTACT 1218
OY	21 LeuAlaProGIUGInLeuileLeuileArGLyLSysAsnProGIyGIYTPTPGLuGLy 40
Db	1219 CTCGCCCTCCTGCAGCTGATTTTCAATCCGAATAAAGAACCCAGTGAGTGTGGAGAAGA 1278
OY	41 GIuLeuGlnAlaArGLyLyLSysArgGlnIlleglYTrPheProAlaSerTYrrVallys 60
Db	1279 GAGCTGCACACGACGTGGAAAAAGCGCCACAATGAGGTGCTTCCAGCTAATTATGTAAAG 1338
OY	61 LeuLeuSerPro 64
Db	1339 CTTCATAAGCCCT 1350
RESULT 3	
LOCUS	AF132672 3812 bp mRNA linear ROD 28-JUN-1999
DEFINITION	Rattus norvegicus EH-domain/SH3-domain containing protein mRNA,
ACCESSION	AF132672
VERSION	AF132672.1 GI:4838525

Pred. No.:	9.5e-33	Length:	4321
Score:	341.00	Matches:	64
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0
US-09-720-934-2_COPY_999_1062 (1-64)	x HSN806384 (1-4321)		
Qy	1	GIgylugluuilealaglnvalillealaserlyrthrplalatrnglyprogluglnleuthr	20
Db	2562	GGAGAAAGAAATTTGGCCAGATTTATTTGGCTTCACACCGCCAGCCGCCGAGACAGCTCACT	2621
Qy	21	leualaproglyglnleuileleuilearghsylvsaaenproglyglYtrpTrpGlugly	40
Db	2622	CTCCGCCCTGGTCAAGCTGATTTTATTCGGAAAAAAGAACCCAGGTGATGTGGGAGGA	2681
Qy	41	GIuIeuglnalaaRGlylyseIysaRGlnilegIytrpPheProAlAAsnTyValIys	60
Db	2682	GAGCTGCAGACACGCTGGGAAAAAGCGCCAGATAGGCTGTTCACGACTAATATGTAAGA	2741
Qy	61	leuIeuSerPro 64	
Db	2742	CTTCTAAGCCT 2753	
RESULT 5			
BD205033	5199 bp DNA linear PAT 17-JUL-2003		
LOCUS			
DEFINITION	Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.		
ACCESSION	BD205033		
VERSION	BD205033.1 GI:33014803		
KEYWORDS	JP 2002511267-A/1.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Embryoto; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 5199)		
AUTHORS	Korenberg,J.R. and Chen,X.N.		
TTITLE	Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof		
JOURNAL	Patent: JP 2002511267-A 1 16-APR-2002;		
COMMENT	CEDARS SINAI HEALTH SYSTEM ET AL		
OS	Homo sapiens (human)		
PN	JP 2002511267-A/1		
PD	16-APR-2002		
PF	16-APR-1999 JP 2000543610		
PR	16-APR-1998 US 60/082007		
PI	JULIE R KORENBERG, XIAO NING CHEN		
PC	C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,		
PC	C12Q1/68,		
PC	G01N33/68//A61K48/00,C12N15/00,C12N5/00		
CC	Isolated SH3 gene relating to myeloproliferative disorders and leukemia		
CC	leukemia		
CC	and utilization thereof.		
FT	Key Location/Qualifiers		
FT	source 1..5199		
FT	Location/Qualifiers		
FEATURES	/organism="Homo sapiens (human)".		
source	1..5199		
	/organism="Homo sapiens"		
	/mol_type="Genomic DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.16e-32	Length:	5199
Score:	341.00	Matches:	64
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-720-934-2_COPY_999_1062 (1-64) x BD205033 (1-5199)
 QY 1 Glycylglutlilealaglnvalilealserythralthrarglyprogluglnleuthr 20
 Db 3202 GGAAAGAGAAATGGCCAGGTTATGGCTCCACACCGCCACCGGCCCGGAGAGCTCACT 3261
 QY 21 Leuylaproglyglnleuileuileuileargylsyrlysasnproglygltyrptptglnlyg 40
 Db 3262 CTCGCCCTGGCTCAGCTGATTTTATATCCGAAAAAGAACCCAGGTGATGGTGGAAAGCA 3321
 QY 41 GluileuglnlaaargyltlysylsyrarglnileglytyrphpepcoalaasnyrVallys 60
 Db 3322 GAGCTGCAGACGACTGGGAAAAAAGCGCAGATAGCTGGTTCCAGCTAATATGTAAG 3381
 QY 61 LeuleuSerPro 64
 Db 3382 CTTCTAAGCCCT 3393
 RESULT 6
 AF127798
 LOCUS DEFINITION Rattus norvegicus EH- and SH3-domain containing protein EHS1 mRNA, complete cds.
 AF127798
 ACCESSION AF127798.1 GI:4835852
 KEYWORDS
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 REFERENCE 1 (bases 1 to 4025)
 AUTHORS Okamoto,M., Schoch,S. and Sudhof,T.C.
 TITLE EHS1/intersectin, a protein that contains EH and SH3 domains and binds to dynamin and SNAP-25. A protein connection between exocytosis and endocytosis?
 JOURNAL J. Biol. Chem. 274 (26), 18446-18454 (1999)
 MEDLINE 99303609
 PUBMED 10373452
 REFERENCE 2 (bases 1 to 4025)
 AUTHORS Okamoto,M., Schoch,S. and Sudhof,T.C.
 TITLE Direct Submission
 JOURNAL Submitted (11-FEB-1999) Center for Basic Neuroscience and HHMI, Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75235, USA
 FEATURES
 source
 location/qualifiers
 1..4025 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 41..3694 /codon_start=1
 /product="EH- and SH3-domain containing protein EHS1"
 /protein_id="AAD0271.1"
 /db_xref="GI:4835853"
 /translation="MAQFPPTPGSGSLDIMAIVTEBRAKHQOQFOSLKPISGFLTPDQA
 RNPFOSGLPOPVLAQIOWALADNMKDMROVRSIMKILIKLQGYOLPPLPWA
 KOOPAAISSAPARGISGMAGMPTTANVAPPMWSIPIVNGSPILVSIPAAPLPLAN
 GAPAVIOPLPFAFAPATIRKSSFSFSSGGSQNTLQAKQSDIVASAPAAEMAVP
 QSSSLKQRLFNHSDKTMGSLTGPQARTIMOSLSLQALASLWNSDIDQGLTA
 BEFLIAHMLIDVMSGQPLPPVLPPEVTPFSFRVRSQSMYSISSAQRLEPERS
 SEDPOQVEKKLPVTFEDKKRKNFPRGMLILEKROALDEQKQERLQALEPABER
 KERROERKROLEIKOLEKORLEORREBERKLEIRERAKRELEROLEWBER
 NRREILLTORNKQEGIVLAKRKLELEFLALNKKQLDEGLDIDRLALNOROE
 IESTNKRERLRIAEITHLQOOLQDSQOMGRILPEKILSDQLKQYQNSLHSDSLT
 LKRALERKELARQLEQLEVEKERSTKQETIDVENNNQLEIRLHSKQLEKQST
 EABLKLEKQERKSLLEKQERKQERQVRRDQWQEHVQEQERKPEBDKLR
 EDSVKKKEERAPAEVDKOSRLFPHQBPAPKAPQAPWPTTEKGLTISAQSAKVVA
 YVALVYFESRSHDEITQGDQIVMVGWEVDSQGEPMGLGGEKGLGWPAAYAA
 EKIPENITPPAPKPYDILTAPAPKALARETAPAPVPTSGSPSTPMNMAVDSQWS
 STNEKPEITMDWDTMAQPSLTVPSAQGLRQSAFTPLATPAPATPVLQGEKEGLO
 AQLVPRAPKDHMLFNKNSDVIIVLEQDMWFGVEVQGGKMGFPKSVVLISPVKR

ORIGIN

STSIDTGPTEAPSSILKRVASPAKPAIPGEFEVAMTYESESHEGLJTFQGHVIVYTK
KDDDMWTGVTGTSIVPRSNVYVILKDSSEGTGKXGSLKEKEFIQAVIATSTATPE
QULIAPGQILIRKKNPGWMBELQMRKROIGMPANVYVILSPGISKITPTELP
KTAVQPAVCQVIGMIDYTAQNDLAFSKQIINVLKSDPDWVKGEVSGQVGLFPSN
YVLTITDMDPSQQ"

Alignment Scores:

Pred. No.: 2,13e-31 Length: 4025
Score: 330.00 Matches: 62
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.77% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x AF127798 (1-4025)

QY 3 GlnlAlaGlnValIleAlaSerYrThAlaThGlyProGluGlnLeuThrLeuAla 22
Db 3260 GAATGCGCCAGGCTATTGCTTCTTCTACACTGCTACGAGGCTCTTAACAGCTCAGCTGAGCT 3319
QY 23 ProGlyGlnLeuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 42
Db 3320 CCGTGTGAGCTGATTTTGTATCCGAAAAAGAACCCCGGTGATGTCGGAAGGAACTA 3379
QY 43 GlnAlaArgGlyIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 62
Db 3380 CAAGCTCGAGGAAAAAGCCGACAGATGAGGTGTTCACGCAAAATATGCAAACTTCTA 3439
QY 63 SerPro 64
Db 3440 AGCCCT 3445

RESULT 7
LOCUS BD205035 5195 bp DNA linear PAT 17-JUN-2003
DEFINITION Isolated SH3 gene relating to myeloproliferative disorders and
leukemia and utilization thereof.

ACCESSION BD205035
VERSION BD205035.1 GI:33014805
KEYWORDS JP 2002511267-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 5195)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Korenberg, J.R. and Chen, X.N.
TITLE Isolated SH3 gene relating to myeloproliferative disorders and
leukemia and utilization thereof
JOURNAL Patent: JP 2002511267-A 3 16-APR-2002;
COMMENT CEDARS SINAI HEALTH SYSTEM ET AL
OS Homo sapiens (human)
PN JP 2002511267-A/3
PD 16-APR-2002
PF 16-APR-1999 JP 2000543610
PR 16-APR-1998 US 60/082007
PI JUIIE R KORENBERG, XIAO NING CHEN
PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,
PC C12Q1/68,
PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and

CC leukemia
CC and utilization thereof.
FH Location/Qualifiers
FT source 1..5195
FT /organism='Homo sapiens (human)'.
Location/Qualifiers

FEATURES
source 1..5195
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 2,81e-31 Length: 5195
Score: 330.00 Matches: 62
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.77% Indels: 0
DB: 6 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x BD205035 (1-5195)

QY 3 GlnlAlaGlnValIleAlaSerYrThAlaThGlyProGluGlnLeuThrLeuAla 22
Db 3452 GAATGCGCCAGGCTATTGCTTCTTCTACACTGCTACGAGGCTCTTAACAGCTCAGCTGAGCTG 3511
QY 23 ProGlyGlnLeuIleuIleuIleuIleuIleuIleuIleuIleuIleu 42
Db 3512 CCGTGTGAGCTGATTTTGTATCCGAAAAAGAACCCCGGTGATGTCGGAAGGAACTG 3571
QY 43 GlnAlaArgGlyIleuIleuIleuIleuIleuIleuIleuIleuIleu 62
Db 3572 CAAGCAGTGGGAAAAAGCCGACAGATGAGGTGTTCACGCTATATGTAAGCTTCTA 3631
QY 63 SerPro 64
Db 3632 AGCCCT 3637

RESULT 8
LOCUS AF114488 5381 bp mRNA linear PRI 16-JUN-2002
DEFINITION Homo sapiens intersecin short isoform (ITSN) mRNA, complete cds.
ACCESSION AF114488
VERSION AF114488.1 GI:4808824
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 5381)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Pucharcos, C., Fuentes, J.J., Casas, C., de la Luna, S., Alcantara, S.,
Arbones, M.L., Soriano, E., Estivill, X. and Pritchard, M.
TITLE Multivalent binding protein expressed in proliferating and
differentiating neurons and overexpressed in Down syndrome
JOURNAL Eur. J. Hum. Genet. 7 (6), 704-712 (1999)

JOURNAL MEDLINE 10482960
PUBMED 99415290
REFERENCE 2 (bases 1 to 5381)
Pucharcos, C., Fuentes, J.J., Pritchard, M. and Estivill, X.
AUTHORS Direct Submission
TITLE Submitted (16-DEC-1998) Medical and Molecular Genetics Center,
Cancer Research Institute, L'Hopitalat de l'lo., Avia.
JOURNAL Castelldefels km. 2,7, Barcelona 08907, Spain
Location/Qualifiers

FEATURES
source 1..5381
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
1..5381
/gene="ITSN"
/gene="ITSN"
269..3931
/gene="ITSN"
/codon_start=1
/product="intersecin short isoform"
/protein_id="MAD29953.1"
/db_xref="GI:4808825"
/translation="MAQFTPPGGSLDIAIVTEERAKHDQFHSCKPISGRTGQQA
RNFPGSGLPQVLAQIWAALADMMNDGRMDQVEFSIAMKLIRKIQGYQSLPSPVPM
KOQPAVIASSAPAFMGIGIASMPPLTAIVAPGSGIPVVGMSPTLVSSVPTAAVPIAN
GAPVIOPLPAFAHPAATLPSKSSFSRSGSGSLNTRLOKQGSFDVAVSPVAVEMAVP
GSSRLKYRQLFNSHDKTMSGHLTGQARTILMOSSLPQQLASISIMWLSIDIDQCKLTA


```
/gene="ITSN"
269..5434
/gene="ITSN"
/codon_start=1
/product="intersectin long isoform"
/protein_id="A029952.1"
/db_xref="GI:4808823"
/translation="MAQPPPPFGSGSLDINALTYERAKHDQGFHSIKPISGFTTGOA
RNFQSGLEPQVLAQIMALADNNDGRNDQVETSLAKLIKLTQSGYQPKLPLPVM
KQPPVATSSAPAFMGQIASMPILTAVPAPMGISLVVMGSPVLVSVPFAAPLAN
GAPVIOPLPAFAPATLIPKSSFSRSGPGLNLTOKAGSFDAVAPVMAFAP
OSSRLKROLFNHDKTMGCHLTGPQARTIMOSSLPQALATNMLSIDODGKTLA
EERILAMHLIDVMSGOLPEVLPPEVTPSPFRRVSGSGISYSTSTGQDPEEVP
LEBDQOLEKPLPVTEDKKRNFPERGNILERRQALLBQKSEBERLAQERAOE
RERERQOEKROLEBERQERLEKQERLEERERKELEERRAARELEERQOME
RNRORLINOQNEODIYVLKAKKTLERLEALNDKQLEGGKODIRCRILTRRO
ELESNKSERLIAETHTLOOLOSQOMGRIPKQILINDQKQONSLHRDGLV
TLKRALEAKELARQIDOLDEYERKTRSLQGEIDIPNNOLKELEIHHKQLOKXS
MEARLKQOEORKEITILEKQKEAORAEORCKMLEHYOQEDERQRLKHEEKL
KRESVYKKGDEKQOEADKGRLEFQOEPAKPAVQAPMTAEKGRPLTSAQENV
KVVYALYPPESRSHDEITIQGDIWMVGWEDSQTGEPMGLGELKGTGMPA
NVAEKIPENEVPAPVPVDTSTAPAPKALRETAPLAVTSSEPTSPNMDAFEST
WPTSTNEKETNDWMAAOPSLTVPSAGOLRORSAFTPATATSSPSPVLGGSEVE
GLOAOLYMPBAKNDHNLNPKNDVITVLEOQMMWGEVQGGKMPKSVYKLISGP
IKRSTMSGSSSPASLKRVAAPKPVSGEFTAMTYTSSBOGDI.TPOCGDYL
VIRKQSDMTGTGYGDKAGVPPSVYVYKDSBSGCTGKTSUGKRELIQVASTAT
GPEQLTLAEGLILIRKKNPGWMBELQARGKROIGWFPANYVKLISPGTITPT
EPKSTALAAVCQVIMDYTAQNDDELAFNKQIINVINKEDPMWKBEVNGQVLF
PSNVYKLTDMPSQOMCSDLHLDMLPTERKQGYIHLLIYENYVNDQLVTEI
FOKLESELTEKEVAMI.FVMKELIMCNIKILKRVKMKSGEMVPMJDLIS
AOLPHMOPYIRPESQNLGANAALIOCKTDEAPDKPERKRAMPRCKMPLSFTIKP
MORATRYPLIKILIENTPBNPDHSHLHAKELKABELCSQNEGVREKENSDELMT
QAHVQEGSELEQVFNSTVNCJGPRKPLKSKYKAKSNELGFLFNFPLLTQITK
PLSSGSTDVFSFKSNLYQMYKTPILFNEVLYKPLTPDSGDEPIFIHSHIVYLR
ASINSRTAMVQKIKAASELYLETEKKEKALVRSORATIGRLMVMVGEITELKP
CSHGKSNPYCEVTWMSOCHITKTIDTIDNPKMNSCOPFIRDLBOEVCLTYEBEDQ
FSPDPLGSTEIRVADIKKQDQSKGPVTKCLLHEVPTGEIVRDLQLDEDE"
```

ORIGIN

Alignment Scores:

Pred. No.:	3..55e-31	Length:	6439
Score:	330.00	Matches:	62
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	96.77%	Indels:	0
DB:	9	Gaps:	0

US-09-720-934-2_COPY_999_1062 (1-64) x AF114487 (1-6439)

```
QY      3  GlnuLlEAlaGlnValIleAlaSerTYrThrAlaThGlyProGluGlnLeuThrIleuAla 22
      |||
Db      3497  GAAATTGCCAGGTTATTGGCTCATACACGCCACCGGCCCGCAGACGCTCACTTCGCC 3556

QY      23  ProGluGlnLeuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 42
      |||
Db      3557  CCGGTCACGCTGATTTTGCATCCAAAAAGAACACCCAGGTGATGCGGGAAGAGACGCTG 3616

QY      43  GlnAlaArgGlyLysLysArgGlnIleGlyTyrPheProAlaAsnTyrValLysLeuLeu 62
      |||
Db      3617  CAAGCAGCGTGGAAAAAGCGCCAGATAGGCTGCTGCCACGTAATTATGTAAAGCTTCTTA 3676

QY      63  SerPro 64
      |||
Db      3677  AGCCCT 3682
```

RESULT 11

AF180522

LOCUS AF180522 1996 bp mRNA linear PRI 05-SEP-1999

DEFINITION Homo sapiens intersectin short form 2 (ITSN) mRNA, partial cds.

ACCESSION AF180522

VERSION AF180522.1 GI:5823551

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1996)

AUTHORS Tsyba,L.O., Kvasha,S.M., Skripkina,I.Y., Anoprienko,O.V., Slavov,D., Tassone,F., Rynditch,A.V. and Gardiner,X.

TITLE Mouse homolog of human chromosome 21 genes

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1996)

AUTHORS Tsyba,L.O., Kvasha,S.M., Skripkina,I.Y., Anoprienko,O.V., Slavov,D., Tassone,F., Rynditch,A.V. and Gardiner,X.

TITLE Direct Submission

JOURNAL Submitted (25-AUG-1999) Department of Molecular Oncogenetics, Institute of Molecular Biology and Genetics of National Academy of Sciences of Ukraine, 150, Zabolotnogo str., Kiev 252627, Ukraine

FEATURES

source

1..1996

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q22.1-q22.2"

/tissue_type="brain"

/dev_stage="infant"

<1..1996

/gene="ITSN"

<1..566

/codon_start=3

/product="intersectin short form 2"

/protein_id="A029952.1"

/db_xref="GI:5823552"

/translation="KSYVYKLISGPRIKSTMSGSSSPASLKRVAAPKPVSGEE

IQAIVSYAPAGEQLTLAPGQILIRKKNPGWMBELQARGKROIGWFPANYVKL

LSPGKSTIPTEPKSTALAAVCQVIMDYTAQNDDELAFNKQIINVINKEDPMW

KSEVNGQVGLPSPNVYKLTDMPSQO"

ORIGIN

Alignment Scores:

Pred. No.:	1.79e-31	Length:	1996
Score:	328.00	Matches:	62
Percent Similarity:	96.88%	Conservative:	0
Best Local Similarity:	96.88%	Mismatches:	2
Query Match:	96.19%	Indels:	0
DB:	9	Gaps:	0

US-09-720-934-2_COPY_999_1062 (1-64) x AF180522 (1-1996)

```
QY      1  GlyGluGlnuLlEAlaGlnValIleAlaSerTYrThrAlaThGlyProGluGlnLeuThr 20
      |||
Db      126  GGAAGAAAGAAATTGCCAGGTTATTGCTCATACACGCCCGGCCCGCAGACGCTCACT 185

QY      21  LeuAlaProGluGlnLeuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 40
      |||
Db      186  CTCGCCCTCGTACGCTGATTTTGCATCCAAAAAGAACACCCAGTGGATGCTGGAAGCA 245

QY      41  GlnLeuGlnAlaArgGlyLysLysArgGlnIleGlyTyrPheProAlaAsnTyrValLys 60
      |||
Db      246  GAGCTCAAGCAGCGTGGAAAAAGCGCCAGATAGGCTGCTGCCACGTAATTATGTAAAG 305

QY      61  LeuLeuSerPro 64
      |||
Db      306  CTTCTAAGCCCT 317
```

RESULT 12

AF064243

LOCUS AF064243 5287 bp mRNA linear PRI 21-NOV-1998

DEFINITION Homo sapiens intersectin short form mRNA, complete cds.

ACCESSION AF064243

VERSION AF064243.1 GI:3859852

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

[illegible]

```

Query Match: 95.89% Indels: 0
DB: 9 Gaps: 0
US-09-720-934-2_COPY_999_1062 (1-64) x AF064243 (1-5287)

OY 3 Glu1le1aagiVal1le1a1aSerTyrThrAlaThrGlyProGluGlnLeuThrLeuAla 22
Db 3335 GAATTGGCCAGGTTATTGCTCTATACACGCCACCGGCCCGGAGAGCTCATCTCGCC 3394
OY 23 ProGlyGlnLeuLeuLeuLeuLeuArgLysLeuAsnProGlyGlyTPTPTGlnGlnGlnLeu 42
Db 3395 CCTGTGAGCTGATTGATTTGATCCGAAAAAGAACCGAGTGATGATGAGGAGAGCTG 3454
OY 43 Glu1laaArgGlyLysLysArgGln1leGlyTrrPheProAlaAsnTyrValLysLeuLeu 62
Db 3455 CAACGACGTGGGAAAAAGCCGCAATAGCGTGTCTCCAGCTAATTATGTAAGCTTCTA 3514
OY 63 SerPro 64
Db 3515 AACCT 3520

RESULT 13
LOCUS AF064244 7247 bp mRNA linear PRI 21-NOV-1998
AF064244
DEFINITION Homo sapiens intersectin long form mRNA, complete cds.
ACCESSION AF064244
VERSION AF064244.1 GI:3859854
KEYWORDS
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 7247)
Guipponi,M., Scott,H.S., Chen,H., Schebesta,A., Rossier,C. and
Antonarakis,S.E.
Two isoforms of a human intersectin (ITSN) protein are produced by
brain-specific alternative splicing in a stop codon
Genomics 53 (3), 369-376 (1998)
99017974
JOURNAL MEDLINE
PUBMED 9793604
2 (bases 1 to 7247)
Guipponi,M., Scott,H.S., Chen,H., Schebesta,A., Rossier,C. and
Antonarakis,S.E.
Direct Submission
Submitted (05-MAY-1998) Genetics and Microbiology, CMU, 1 rue
Michel-Servet, Geneva, 4 CH-1211, Switzerland
Location/Qualifiers
1..7247
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
/tissue_type="brain"
/dev_stage="fetus"
107..5272
/codon_start=1
/product="intersectin long form"
/protein_id="AAC78611.1"
/db_xref="GI:3859855"
/translation="MAQPTPTFGSGSLDIMAIVTERAKHQDFSLNPISGFTTGDA)
RNFPGSLPOPVLAIQIMALADMNNDRMQOVESIMAKIILKLCYQYASALPVM
KQDFVALSSAPRPMGMSIAGMPILTNVAQMSIPVAGSSVTPRAAPVPLANN
GAPVIOPLPAPHPAATLTKSSSFFSGSGSGLNLTOKAASPDVASVPPVMAVAP
OSSRIKROLPNSHDKTMSGHLNTPKAPKATILMOSLSPQALASINWSDIDOGKLTAA
EEFLAHNLIDVMSGQPLPVPVPEPTIPSPFRVKSGLSVISSTSVQRLPPEPV
LEBQOQLEKKLPTVFEDEKRENFERNGLLEKRRQALLIQKREDERLAQLEBAQDE
RKREBOEQERKRQLELEKQELERQELERQEBERKEIERRAKXELRQRLQLEMBE
RNSRQELNQRKEQED1VILAKQOQLELELELAKNKKQLEBKQODICRLTTCQEQ
EISTNSRELRLAIEITHLQOQOESQOMGRILPEKQIINDQLOQVQNSIHRSDIVQ
TKRALEAKELARQHLDDQIDVEYKETRSLQETIDFNNOIKELREIHNKQLOQKQS
MEARLQKQEQERKTIIELEKQKEAQRARERBKQMLEHYQOEENHRPKIHEEKL
KRESVKKQKQGEERKQEQADKIGRLPHQEOEPKAPVQAPWSTAEGPLTISKQENV

```

```

KVVYRYALYPEFSRSHDEITIOEGDIWMKGEWVDSOTGEPGLGELGKTGMPPA
NYAEKIPENVEVPAVPKPVDTSTSAAPKALAREPAPLATSSBPTTPNNMDFST
WPTSTNKEPTDMDWADAAPSLTIVEAGOLRORSATPTATNSSSPVCGCEKYE
GLQOALYPRKAKDNHNRKNDVITVLQODMWFGEVQGMPPKSVLQISLP
IRKSTMSDSSSPASLKRVASPAAPVVGSEFLAMYBESQODLTFQOQDVL
VTKKDGMWTVGDKAGVFPNNYVRKDSGSGTAKTSLSGKPEQDLAVIASYAT
GPEQLTAPGOLLIRKKNPGMWEGLORGGKROIGMPANVVKILNKGTSKPTT
EPPESTLAACVQIGMYDYTAONDDLAENKQOILNVLKEDPDMWKGWNOYGF
PSNVKLTMDPSQWCSLDLHMLTPTERRQGITHELIYTERYVNDLQUTBI
FOYPMSESLTTEREVAMITVNMKELTMCNLIKALKARVKKMSGEMPKMTDILS
AQLPHMQPIYRFSRQNGAALIQOKTDEAPDEKEFVRLMEPRCKGMLSPILP
MOVTRYPLIKNILENTPENHDSHSLKALKAELCSQVNEGYEAKNSRLMEI
QAHVOCEGLSEOLFVNSVNCIGPKRKLHSGKLYKAKNKELYGLPNDLLTQITF
PLSGSGDKVSPSKSNLOVMYKTPYFLNVLVKLPDPSGDEPIPHISHIDYVTLR
AESINERTAVQKIKASSEIYITEKKRKRAIVSQRATGIGRLHNVSEIELEK
CRSHGKSNPYCEVWMSQCHITITLQUTLNPKNNSNCFIRLREDEVLCITVFEHQ
FSPDPLGRTEIRVADIKKQGSKPGVKTLHHEVPTGEIVRLDQLDFEB"
misc_feature
167..406
/note="encodes EH domain"
misc_feature
767..1936
/note="encodes EH domain"
misc_feature
2324..2524
/note="encodes SH3 domain"
misc_feature
2843..3019
/note="encodes SH3 domain"
misc_feature
3110..3286
/note="encodes SH3 domain"
misc_feature
3326..3520
/note="encodes SH3 domain"
misc_feature
3569..3748
/note="encodes SH3 domain"
misc_feature
3836..4390
/note="encodes SH3 domain"
misc_feature
4649..4819
/note="encodes GDS domain"
misc_feature
4895..5143
/note="encodes PH domain"
misc_feature
/note="encodes C2 domain"
ORIGIN
Alignment Scores:
Pred. No.: 9.61e-31 Length: 7247
Score: 327.00 Matches: 61
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.39% Mismatches: 0
Query Match: 95.89% Indels: 0
Gaps: 0
DB: 9
US-09-720-934-2_COPY_999_1062 (1-64) x AF064244 (1-7247)
QY 3 GlnlAlaGlnValIleAlaSerYrThrAlaThrglyProgluGlnleuThrleuAla 22
Db 3335 GAATATGCCAGGTATATGCTCATACACCGCCAGCGCCCGCAGAGCTCACTCTCGCC 3394
QY 23 ProglYglnleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 42
Db 3395 CCGGTCACTGATTTTGAATCCGAAAAAGAACCCAGTGGATGGTGGAGAGAGCTG 3454
QY 43 GlnAlaArgGlylyslYsArgGlnIleGlyTrrPheProAlaAsnTyVallyslLeuLeu 62
Db 3455 CAAGCACGTGGGAAAAAGCGCCAGATAGGCTGCTCCAGCTAATTATGTAAGCTTCTA 3514
QY 63 SerPro 64
Db 3515 AACCT 3520
RESULT 14
AF169621 1133 bp mRNA linear ROD 12-AUG-1999
LOCUS AF169621
DEFINITION Mus musculus Esel protein mRNA, partial cds.
ACCESSION AF169621
VERSION AF169621.1 GI:5731280
KEYWORDS
SOURCE Mus musculus (house mouse)

```

```

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1133)
Teyba,L.O., Kvasha,S.M., Skripkina,I.Y., Anoprienko,O.V.,
Slavov,D., Tassone,F., Rynditch,A.V. and Gardiner,K.
TITLE
Mouse homologues of human chromosome 21 genes
JOURNAL
Unpublished
2 (bases 1 to 1133)
Teyba,L.O., Kvasha,S.M., Skripkina,I.Y., Anoprienko,O.V.,
Slavov,D., Tassone,F., Rynditch,A.V. and Gardiner,K.
TITLE
Direct Submission
JOURNAL
Submitted (15-JUL-1999) Department of Molecular Oncogenetics,
Institute of Molecular Biology and Genetics of National Academy of
Sciences of Ukraine, 150, Zabolotnogo str., Kiev 252627, Ukraine
FEATURES
source
1..1133
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/dev_stage="15-day embryo"
<1..749
/note="similar to Homo sapiens Intersectin"
/codon_start=3
/product="Esel protein"
/protein_id="AAD4848.1"
/db_xref="GI:5731281"
/translation="VRKSTSIDGPTEPSASLKRVASPAKPAIPGEFFIATYESS
EGDLPFOQGDVIVTKDGMWTVGDKAGVFPNNYVRKDSGSGTAKTSLSGKPEQDLAVIASYAT
GPEQLTAPGOLLIRKKNPGMWEGLORGGKROIGMPANVVKILNKGTSKPTT
EPPESTLAACVQIGMYDYTAONDDLAENKQOILNVLKEDPDMWKGWNOYGF
PSNVKLTMDPSQWCSLDLHMLTPTERRQGITHELIYTERYVNDLQUTBI
FOYPMSESLTTEREVAMITVNMKELTMCNLIKALKARVKKMSGEMPKMTDILS
AQLPHMQPIYRFSRQNGAALIQOKTDEAPDEKEFVRLMEPRCKGMLSPILP
MOVTRYPLIKNILENTPENHDSHSLKALKAELCSQVNEGYEAKNSRLMEI
QAHVOCEGLSEOLFVNSVNCIGPKRKLHSGKLYKAKNKELYGLPNDLLTQITF
PLSGSGDKVSPSKSNLOVMYKTPYFLNVLVKLPDPSGDEPIPHISHIDYVTLR
AESINERTAVQKIKASSEIYITEKKRKRAIVSQRATGIGRLHNVSEIELEK
CRSHGKSNPYCEVWMSQCHITITLQUTLNPKNNSNCFIRLREDEVLCITVFEHQ
FSPDPLGRTEIRVADIKKQGSKPGVKTLHHEVPTGEIVRLDQLDFEB"
ORIGIN
Alignment Scores:
Pred. No.: 2.31e-31 Length: 1133
Score: 325.00 Matches: 61
Percent Similarity: 98.39% Conservative: 0
Best Local Similarity: 98.39% Mismatches: 1
Query Match: 95.31% Indels: 0
Gaps: 0
DB: 10
US-09-720-934-2_COPY_999_1062 (1-64) x AF169621 (1-1133)
QY 3 GlnlAlaGlnValIleAlaSerYrThrAlaThrglyProgluGlnleuThrleuAla 22
Db 3315 GAATATGCCAGGTATATGCTCATACCGCTCACTGCTCCGAAACAACCTCGCT 374
QY 23 ProglYglnleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 42
Db 375 CCGGTCACTGATTTTGAATCCGAAAAAGAACCCAGTGGATGGTGGAGAGAACTG 434
QY 43 GlnAlaArgGlylyslYsArgGlnIleGlyTrrPheProAlaAsnTyVallyslLeuLeu 62
Db 435 CAAGCTCGAGGAAAAAGCGCCAGATAGGCTGCTCCAGCAATTATGTCACAACTTCTA 494
QY 63 SerPro 64
Db 495 AGCCCC 500
RESULT 15
AF132478 3723 bp mRNA linear ROD 09-MAR-1999
LOCUS AF132478
DEFINITION Mus musculus Esel protein mRNA, complete cds.
ACCESSION AF132478
VERSION AF132478.1 GI:4378884
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 3723)
Sengar,A.S., Wang,W., Bishay,J., Cohen,S. and Egan,S.E.
AUTHORS

```

TITLE

The EH and SH3 domain Eps proteins regulate endocytosis by linking to dynamin and Eps15

Job time : 1469.86 secs

JOURNAL

EMBO J. 18 (5), 1159-1171 (1999)

MEDLINE

99164083

PUBMED

10064583

REFERENCE

2 (bases 1 to 3723)

AUTHORS

Sengar, A.S., Wang, W., Cohen, S., Bishay, J. and Egan, S.E.

TITLE

Submitted (02-MAR-1999) Programs in Cancer & Blood

JOURNAL

Research/Developmental Biology, The Hospital for Sick Children, 555

FEATURES

Location/Qualifiers

SOURCE

1. .3723

CDS

/organism="Mus musculus"

FEATURES

/mol_type="mRNA"

SOURCE

/db_xref="taxon:10090"

FEATURES

1. .3642

SOURCE

/codon_start=1

FEATURES

/product="Eps1 protein"

SOURCE

/protein_id="ABD19746.1"

FEATURES

/db_xref="GI:4378885"

SOURCE

/translation="MAQPTFGSLDVVAITVEBRAXHDQFLSLKFIAGFTTGDQA

FEATURES

RNFFQSGLPQVLAQIWMALADNNDRMDQVEFSIAMKLIKIKQYQLPSTLPVM

SOURCE

KQCPVAISSAPFAGIGIASMPPLTAAPVPMGSIPIVGMSPPLVSVPPAAPPPLAN

FEATURES

GAPVIOPLPAPAPAPATWPKSSFSRSGSOLNTKLQKASFDVASAPPAAWAVP

SOURCE

OSRLKRYQLFNSHDKTMSGHLTGPOARTIMOSSLPQALASIMNLSIDIDQDKLTA

FEATURES

EEFLAMHLIVNMSGQPLPVLPPPEYIPPSFRYRSGSGMSYISSSVDRPLPEPS

SOURCE

SEDEQPEPKLPVTFEDKRENFEKSVLEKROQLLEQKQERLAQLERABQER

FEATURES

KERROEQEAKROLEKOLKRELEKREBERKELEBRERLEKRELEKRELEKRELE

SOURCE

NRROELNORNEOBGVVLEKARKELEFELELNDKQLEGLKODIRCLATQROE

FEATURES

LESTNKSRELELAEITHLQOOLESQWLGRLPEKQILSDOLKQVQNSLHPSLILT

SOURCE

LKRLAEAKELARCOLRBDLEVERETRSKLEIDVFNNOIKELREIHSKQQLOKORSI

FEATURES

EAARLKQERKRSLELEKQKDAORRVOEDKQLEHVOQEOBPRPRKPEBEDIKR

SOURCE

EDSVKKEAEERAKEMQDKOSRLPHQEPKATQAPWSTTEKPLTISAQSVKV

FEATURES

VYRALYPFESRSHDRTIOPDIWMDESQTEBGMGLGKTKGTMFPANVAKLP

SOURCE

ENEVPTPAKPVTDLTAPAPKALAEETPAPIVTSSEPTTPNNMADFSSTWPSSE

FEATURES

KPETNDMTMAOPELTIVSAGOLRORAPATAGSSPSPVLQOGKEVGLQAOAL

SOURCE

YPMRAKKNHINFNKSDVITVLEODQMMWFGVOGKGMFPKSYVKLISGPVKSTSI

FEATURES

DGTETESPASUKRVAAPRAIPREFIAMTYESSROGDLTFOGQGVIVTKKDD

SOURCE

WMGTGVDKSGVFPENYVRLKDSBSSGTAGTSGISGKKPEIAQVIASTAAATPEQTL

FEATURES

APGDLIRKKNPGMGEGLQARGKRQIGMFPANVVKLSPGTSKITPTELEPTAV

SOURCE

QPAVCOVIGMYDYTAQNDELAFSGQIINVLNKEDPDMKGEVSGQVGLFSPNVKL

FEATURES

TTMDPSQ"

ORIGIN

Alignment Scores:

Pred. No.: 8,366-31 Length: 3723
Score: 325.00 Matches: 61
Percent Similarity: 98.39% Conservative: 0
Best Local Similarity: 98.39% Mismatches: 1
Query Match: 95.31% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x AF132478 (1-3723)

QY 3 GlnlealaglnvallealaserTyrthrAlaThrGlyProGluGlnleuThrleuAla 22

Db 3208 GAAATTGCCAGGTATTGCTTCCCTACGCTGCTACTGCTCCGAACAACCTCACCTGGCT 3267

QY 23 ProGlyGlnleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 42

Db 3268 CCTGGGACGCTGATCTGATCGGAAAAAGAACCCAGGTGATGGGGAAGAGAACTG 3327

QY 43 GlnAlaArgGlyLysLysArgGlnIleGlyTTPheProAlaAsnTyValLysleuLeu 62

Db 3328 CAAGCTCAGGAGAAAGCGCGATAGGTGGTTTCCAGCAATTAATGTCAAACTTCTA 3387

QY 63 SerPro 64

Db 3388 AGCCCC 3393

Search completed: July 1, 2004, 16:15:23

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:03:00 ; Search time 44.815 Seconds
(without alignments)
792.522 Million cell updates/sec

Title: US-09-720-934-2_COPY_999_1062
Sequence: 341
1 GERIAVIVSYATATGPEQLT.....RKKRQIGWFPANYVVLSP 64

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 segs, 277475446 residues
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09120934/rnat_30062004_064540_13455/app_query.fasta_1.1386
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09120934_@CEN_1_1_284_@rnat_30062004_064540_13455 -NCPU=6 -ICPU=3
-NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	253	74.2	747	4	US-08-630-915A-39
2	253	74.2	2873	4	US-08-630-915A-193
3	113.5	33.3	1392	1	US-08-475-894-3
4	113.5	33.3	1392	1	US-08-484-710-3
5	113.5	33.3	1392	2	US-08-484-709-3
6	113.5	33.3	1392	2	US-08-474-697-3
7	113.5	33.3	1659	1	US-08-475-894-1
8	113.5	33.3	1659	1	US-08-484-710-1
9	113.5	33.3	1659	2	US-08-484-709-1
10	113.5	33.3	1659	3	US-08-474-697-1
11	113.5	33.3	3143	4	US-08-671-354-1
12	111	32.6	2757	1	US-08-306-691B-48

13	111	32.6	2757	5	PCT-US93-06251-79	Sequence 79, Appl
14	109	31.0	2793	1	US-07-646-537B-1	Sequence 1, Appl1
15	107.5	32.5	1414	4	US-09-023-655-1356	Sequence 1356, Ap
16	106.5	31.2	971	4	US-08-630-915A-197	Sequence 197, Ap
17	106.5	31.2	1457	3	US-09-444-053-3	Sequence 3, Appl1
18	104	30.5	3111	4	US-10-014-882-1	Sequence 1, Appl1
19	104	30.5	3518	4	US-10-014-882-3	Sequence 3, Appl1
20	102.5	30.1	767	4	US-09-328-475C-155	Sequence 155, App
21	102.5	30.1	1045	1	US-08-475-894-5	Sequence 5, Appl1
22	102.5	30.1	1045	2	US-08-484-710-5	Sequence 5, Appl1
23	102.5	30.1	1045	2	US-08-484-709-5	Sequence 5, Appl1
24	102.5	30.1	1045	3	US-08-474-697-5	Sequence 5, Appl1
25	102.5	30.1	1895	4	US-09-620-312D-397	Sequence 397, App
26	102.5	30.1	3870	1	US-08-138-641-1	Sequence 1, Appl1
27	102.5	30.1	3870	1	US-08-138-133-1	Sequence 1, Appl1
28	102.5	30.1	3893	1	US-08-138-641-3	Sequence 3, Appl1
29	102.5	30.1	3893	1	US-08-138-133-3	Sequence 3, Appl1
30	100.5	29.5	4079	4	US-09-016-434-1247	Sequence 1247, Ap
31	99.5	29.2	970	4	US-09-833-381-1294	Sequence 1294, Ap
32	98	28.7	2547	4	US-09-765-298A-9	Sequence 9, Appl1
33	98	28.7	4079	4	US-09-016-434-1246	Sequence 1246, Ap
34	98	28.7	4242	4	US-09-023-655-1201	Sequence 1201, Ap
35	96.5	28.3	822	4	US-09-328-475C-187	Sequence 187, App
36	96.5	28.3	1024	4	US-09-328-475C-3	Sequence 3, Appl1
37	95.5	28.0	1458	2	US-08-942-423-68	Sequence 68, Appl
38	95.5	28.0	2003	2	US-08-942-423-1	Sequence 1, Appl1
39	95.5	28.0	2003	4	US-08-630-915A-25	Sequence 25, Appl
40	95	27.9	5398	3	US-09-356-952-11	Sequence 11, Appl
41	94	27.6	4762	4	US-09-300-958A-30	Sequence 30, Appl
42	93	27.3	467	4	US-09-300-958A-28	Sequence 28, Appl
43	93	27.3	1382	4	US-09-820-005-1	Sequence 1, Appl1
44	92	27.0	1349	4	US-09-023-655-1176	Sequence 1176, Ap
45	91.5	26.8	3345	4	US-08-630-915A-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-08-630-915A-39
; Sequence 39, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLES, Dana M.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 747 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-630-915A-39

Alignment Scores:
Pred. No.: 1.52e-24
Score: 253.00
Percent Similarity: 87.10%
Best Local Similarity: 75.81%
Query Match: 74.19%
DB: 4

Length: 747
Matches: 47
Conservative: 7
Mismatches: 8
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x US-08-630-915A-39 (1-747)

QY 3 GlnlleaaglnvallealeaSerTyrThrAlaThrGlyProGluGlnLeuThrLeuAla 22
Db 334 GAGATTGCTCAGGTAACTTCAATATGCTTCTGTTCTGAAACAATTAGCCTTGCA 393

QY 23 ProGluGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 42
Db 394 CCAGGACAGTTAAATTATTTCTAAAGAAAATACAACTGGGTGGTGGCAAGAGAGTTA 453

QY 43 GlnlaaRgLyLysLysArgGlnlleGlyTrpPheProAlaAsnTyrValLysLeuLeu 62
Db 454 CAGGCCAGAGAAAAAGCCGACAGAAAGAGATGTTCTCGCAGTCATGTTAACTTTTG 513

QY 63 SerPro 64
Db 514 GGTCCA 519

RESULT 2

US-08-630-915A-193
Sequence 193, Application US/08630915A
Patent No. 6309820

GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLES, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630.915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 193:
SEQUENCE CHARACTERISTICS:
LENGTH: 2873 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-08-630-915A-193

Alignment Scores:
Pred. No.: 8.64e-24
Score: 253.00
Percent Similarity: 87.10%
Best Local Similarity: 75.81%
Query Match: 74.19%
DB: 4

Length: 2873
Matches: 47
Conservative: 7
Mismatches: 8
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x US-08-630-915A-193 (1-2873)

QY 3 GlnlleaaglnvallealeaSerTyrThrAlaThrGlyProGluGlnLeuThrLeuAla 22
Db 1119 GAGATTGCTCAGGTAACTTCAATATGCTTCTGTTCTGAAACAATTAGCCTTGCA 1178

QY 23 ProGluGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 42
Db 1179 CCAGGACAGTTAAATTATTTCTAAAGAAAATACAACTGGGTGGTGGCAAGAGAGTTA 1238

QY 43 GlnlaaRgLyLysLysArgGlnlleGlyTrpPheProAlaAsnTyrValLysLeuLeu 62
Db 1239 CAGGCCAGAGAAAAAGCCGACAGAAAGAGATGTTCTCGCAGTCATGTTAACTTTTG 1298

QY 63 SerPro 64
Db 1299 GGTCCA 1304

RESULT 3

US-08-475-894-3
Sequence 3, Application US/08475894
Patent No. 5641748

GENERAL INFORMATION:
APPLICANT: Yen-Ming Hsu
APPLICANT: THE CAIP-LIKE GENE FAMILY
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475.894
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-191
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1392 base pairs
TYPE: nucleic acid
STRANDEDNESS: both


```
OY      2  GluGluIleAlaGlnValIleAlaSerTyrThrAlaThrGlyProGluGlnIleuThrIeu 21
Db      754 AAGGATTACTGCAAGTAAATATTTCCATATGAGGACAGAAATGATGATGATTCACATC 813
OY      22 AlaProGlyGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 39
Db      814 AAAGAGAGATATATGATCTCATCATTAAGCATCATGACATGACGTAGCTGGTGGGAA 873
OY      40 GlyGluIleuGlnAlaArgGlyLysLysArgGlnIleGlyTyrPheProAlaAsnTyrVal 59
Db      874 GGAGAGCTGAACGGCAGACGA-----GGCGGTTCCTCCGATTAACCTTCGTG 918
OY      60 LysIleuSerPro 64
Db      919 AAGTTACTTCACCG 933

RESULT 6
US-08-474-697-3
; Sequence 3, Application US/08474697
; Patent No. 6171800
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,697
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1392 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-474-697-3

Alignment Scores:
Pred. No.:          9.52e-06          Length:          1392
Score:              113.50           Matches:          25
Percent Similarity: 55.38%           Conservative:     11
Best Local Similarity: 38.46%         Mismatches:      22
Query Match:        33.28%           Indels:          7
Gaps:                3               Gaps:            2

US-09-720-934-2_COPY_999_1062 (1-64) x US-08-474-697-3 (1-1392)
OY      2  GluGluIleAlaGlnValIleAlaSerTyrThrAlaThrGlyProGluGlnIleuThrIeu 21
Db      754 AAGGATTACTGCAAGTAAATATTTCCATATGAGGACAGAAATGATGATGATTCACATC 813
OY      22 AlaProGlyGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 39
Db      814 AAAGAGAGATATATGATCTCATCATTAAGCATCATGACATGACGTAGCTGGTGGGAA 873
OY      60 LysIleuSerPro 64
Db      919 AAGTTACTTCACCG 933
```

```
OY      40 GlyGluIleuGlnAlaArgGlyLysLysArgGlnIleGlyTyrPheProAlaAsnTyrVal 59
Db      874 GGAGAGCTGAACGGCAGACGA-----GGCGGTTCCTCCGATTAACCTTCGTG 918
OY      60 LysIleuSerPro 64
Db      919 AAGTTACTTCACCG 933

RESULT 7
US-08-475-894-1
; Sequence 1, Application US/08475894
; Patent No. 5641748
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,894
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-191
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1659 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-475-894-1

Alignment Scores:
Pred. No.:          1.19e-05          Length:          1659
Score:              113.50           Matches:          25
Percent Similarity: 55.38%           Conservative:     11
Best Local Similarity: 38.46%         Mismatches:      22
Query Match:        33.28%           Indels:          7
Gaps:                1               Gaps:            2

US-09-720-934-2_COPY_999_1062 (1-64) x US-08-475-894-1 (1-1659)
OY      2  GluGluIleAlaGlnValIleAlaSerTyrThrAlaThrGlyProGluGlnIleuThrIeu 21
Db      1021 AAGGATTACTGCAAGTAAATATTTCCATATGAGGACAGAAATGATGATGATTCACATC 1080
OY      22 AlaProGlyGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 39
Db      1081 AAAGAGAGATATATGATCTCATCATTAAGCATCATGACATGACGTAGCTGGTGGGAA 1140
OY      40 GlyGluIleuGlnAlaArgGlyLysLysArgGlnIleGlyTyrPheProAlaAsnTyrVal 59
Db      1141 GGAGAGCTGAACGGCAGACGA-----GGCGGTTCCTCCGATTAACCTTCGTG 1185
OY      60 LysIleuSerPro 64
Db      919 AAGTTACTTCACCG 933
```

```

APPLICANT : Yen-Ming Hsu
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
City: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,709
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-192
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1659 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-484-709-1

Alignment Scores:
Pred. No.:      1.19e-05    Length:      1659
Score:          113.50     Matches:      25
Percent Similarity: 55.38%   Conservative: 11
Best Local Similarity: 38.46% Mismatches:    22
Query Match:      33.28%    Indels:       7
DB:                2        Gaps:           2
US-09-720-934-2_COPY_999_1062 (1-64) x US-08-484-709-1 (1-1659)
QY      2 GluGluLeuIleAlaValIleAlaSerTyrThrAlaThrGlyProGluGlnLeuThrLeu 21
Db      1021 AAGGATTCACGCCAAGAATTAATTTCCATATGAGCGCAAGATGATGATGAATTCACATC 1080
QY      22 AlaProGlyGlnLeuIleLeuIleArgLysAsn-----ProGlyGlyTyrPTrGlu 39
Db      1081 AAAGAAGGAGATATGATCATCTGCATCATTAAGACTGCATGCATGACGTGGTGGA 1140
QY      40 GlyGluLeuGlnAlaArgGlyLysArgGlnIleGlyTyrPheProAlaSerTyrVal 59
Db      1141 GGAGAGCTGAAAGCGACAGCA-----GGCGTGTCCCGATAACTTCGTG 1185
QY      60 LysLeuLeuSerPro 64
Db      1186 AAGTACTTCCACC 1200

RESULT 10
US-08-474-697-1
Sequence 1, Application US/08474697
Patent No. 6171800
GENERAL INFORMATION:
APPLICANT: Yen-Ming Hsu
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
City: Boston
```

STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,697
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BCP-193
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1659 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-474-697-1

Alignment Scores:
Pred. No.: 1,19e-05
Score: 113.50
Percent Similarity: 55.38%
Best Local Similarity: 38.46%
Query Match: 33.28%
DB: 3
Length: 1659
Matches: 25
Conservative: 11
Mismatch: 22
Indels: 7
Gaps: 2

US-09-720-934-2_COPY_999_1062 (1-64) x US-08-474-697-1 (1-1659)

QY 2 GluGluIleAlaGlnValIleAlaSerThrAlaThrGlyProGlnGlnLeuThrLeu 21
Db 1021 AAGGATTACTGCAAGATATATTTCCATATAGAGCAGCAATGATGAAATGACATC 1080
QY 22 AlaProGlyGlnLeuIleLeuIleArgIleLysAsn-----ProGlyGlyTyrTrpGlu 39
Db 1081 AAAGAGAGATATATGACTCTCATCATATAAGACTGACATGACGTAGCGTGGGAA 1140
QY 40 GlyGluLeuGlnAlaArgGlyLysLysArgGlnIleGlyTyrPheProAlaAsnTyrVal 59
Db 1141 GGAGAGCTGAACGGCAGACGA-----GGCGTGTTCCTCCGATTAACCTTCGTG 1185
QY 60 LysLeuLeuSerPro 64
Db 1186 AAGTTACTTCACCG 1200

RESULT 11
US-08-671-354-1
Sequence 1, Application US/08671354
Patent No. 6423824
GENERAL INFORMATION:
APPLICANT: Hsu, Yen-Ming
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95

SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/671,354
FILING DATE: 27-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/486,344
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/484,709
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/475,894
FILING DATE: 07-JAN-1995
APPLICATION NUMBER: US 08/475,710
FILING DATE: 07-JAN-1995
APPLICATION NUMBER: US 08/474,697
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Louis
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: 10274/009005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3143 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 436...2412
US-08-671-354-1

Alignment Scores:
Pred. No.: 2,73e-05
Score: 113.50
Percent Similarity: 55.38%
Best Local Similarity: 38.46%
Query Match: 33.28%
DB: 4
Length: 3143
Matches: 25
Conservative: 11
Mismatch: 22
Indels: 7
Gaps: 2

US-09-720-934-2_COPY_999_1062 (1-64) x US-08-671-354-1 (1-3143)

QY 2 GluGluIleAlaGlnValIleAlaSerThrAlaThrGlyProGlnGlnLeuThrLeu 21
Db 1222 AAGGATTACTGCAAGATATATTTCCATATAGAGCAGCAATGATGAAATGACATC 1281
QY 22 AlaProGlyGlnLeuIleLeuIleArgIleLysAsn-----ProGlyGlyTyrTrpGlu 39
Db 1282 AAAGAGAGATATATGACTCTCATCATATAAGACTGACATGACGTAGCGTGGGAA 1341
QY 40 GlyGluLeuGlnAlaArgGlyLysLysArgGlnIleGlyTyrPheProAlaAsnTyrVal 59
Db 1342 GGAGAGCTGAACGGCAGACGA-----GGCGTGTTCCTCCGATTAACCTTCGTG 1386
QY 60 LysLeuLeuSerPro 64
Db 1387 AAGTTACTTCACCG 1401

RESULT 12
US-08-306-691B-48
Sequence 48, Application US/08306691B
Patent No. 5734039
GENERAL INFORMATION:
APPLICANT: Calabretta, Bruno
APPLICANT: Skorski, Tomasz
TITLE OF INVENTION: ANTISENSE
TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seidel, Gonda, Lavoigna & Monaco, P.C.

```

1 STREET: Two Penn Center, Suite 1800
2 CITY: Philadelphia
3 STATE: Pennsylvania
4 COUNTRY: U.S.A.
5 ZIP: 19102
6
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
9
10 COMPUTER: IBM PS/2
11
12 OPERATING SYSTEM: MS-DOS
13 SOFTWARE: Wordperfect 5.1
14
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER: US/08/306.691B
17 FILING DATE: September 15, 1994
18
19 CLASSIFICATION: 514
20
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER:
23 FILING DATE:
24
25 ATTORNEY/AGENT INFORMATION:
26 NAME: Monaco, Daniel A.
27 REGISTRATION NUMBER: 30,480
28 REFERENCE/DOCKET NUMBER: 8321-8
29
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: (215) 568-8383
32 TELEFAX: (215) 568-5549
33
34 TELEX: No. 5734039e
35
36 INFORMATION FOR SEQ ID NO: 48:
37
38 SEQUENCE CHARACTERISTICS:
39
40 LENGTH: 2757 base pairs
41 TYPE: nucleic acid
42 STRANDEDNESS: double
43
44 TOPOLOGY: linear
45
46 US-08-306-691B-48
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
9
```

```

1      OPERATING SYSTEM:  PC-DOS/MS-DOS
2      SOFTWARE:  Patent In Release #1.0, Version #1.25
3      CURRENT APPLICATION DATA:
4      APPLICATION NUMBER:  8586
5      FILING DATE:  19930630
6      CLASSIFICATION:
7      ATTORNEY/AGENT INFORMATION:
8      NAME:  DiGiullo, Frank S.
9      REGISTRATION NUMBER:  31,346
10     REFERENCE/DOCKET NUMBER:  8586
11     TELECOMMUNICATION INFORMATION:
12     TELEPHONE:  516-742-4343
13     TELEFAX:  516-742-4366
14     TELEX:  230 901 SANS UR
15     INFORMATION FOR SEQ ID NO:  79:
16     SEQUENCE CHARACTERISTICS:
17     LENGTH:  2757 base pairs
18     TYPE:  nucleic acid
19     STRANDEDNESS:  double
20     TOPOLOGY:  linear
21     MOLECULE TYPE:  DNA (genomic)
22     DCT-US93-06251-79

```

Alignment Scores:	
Pred. No.:	4.93e-05
Score:	111.00
Percent Similarity:	58.82%
Best Local Similarity:	41.18%
Query Match:	37.55%
DB:	5
Gaps:	2
Length:	275
Matches:	21
Conservative:	9
Mismatches:	15
Indels:	6
Gaps:	2

US-09-720-934-2_COPY_999_1062 (1-64) x PCT-US93-06251-79 (1-2757)

QY	11	TYTTHRLAHHRGYPRGGLGINTLETHLEALPRGGLGINTLEULE---	Leu11e	29
	..:	..:	..:	..:
Db	2343	TTCGGCCCGGTACCGCTTCAGAGCTTCGCTCAAGAGGATGACATCATCAAGATCTT		2402
QY	30	ARGLYSLYSANPRGGLGINTYTPPTPGGLGINTLEUNALAAARGLYLYSLYARG		49
	..:	..:	..:	..:
Db	2403	AACAGACAGGACACAGCAGGCTGTGGCCGAGGAGATCATATGACCGG-----		2450
	..:	..:	..:	..:

```
Qy      50  GlnIleGlyTPDPheProAlaAsnTyrrValLys 60
      ::|||
Db      2451  ---GTTGGCTGTCTCCCTGCACACTACTGCGAG 2480
```

RESULT 14
US-07-646-537B-1
Sequence 1, Application US/07646537B
Patent No. 5348864
GENERAL INFORMATION:
APPLICANT: Bairdacid, Mariano
TITLE OF INVENTION: Vav Proto-Oncogene Protein
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bristol-Myers Squibb Company
STREET: P.O. Box 4000
CITY: Princeton
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08543-4000
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/646,537B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gaul, Timothy J
REGISTRATION NUMBER: 33,111
REFERENCE/DOCKET NUMBER: DC10

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 921-5901
TELEFAX: (609) 921-4526
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2793 base pairs
type: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 14..2545
US-07-646-537B-1

Alignment Scores:
Pred. No.: 9.21e-05 Length: 2793
Score: 109.00 Matches: 21
Percent Similarity: 58.82% Conservative: 9
Best Local Similarity: 41.18% Mismatches: 15
Query Match: 31.96% Indels: 6
DB: 1 Gaps: 2

US-09-720-934-2_COPY_999_1062 (1-64) X US-07-646-537B-1 (1-2793)
QY 11 TyrThralathrglyProgluInleuthrleuAlaProgluInleuile--Leuile 29
Db 2387 TTCTGTGCGCCGGACAGGTCCGAACGTGTCCTTAAGAGGGGTGCATCATCATCATCCTC 2446
QY 30 ArgylstysaAnProgluGlyTTPTrpInluGlyluGluGlnAlaArgGlyLylsArg 49
Db 2447 AATAAGAGGACACGACGAGGCTGGTGGGCTGGAGATCTACGGCCG----- 2494
QY 50 GlnleGlyTrpPheProAlaAsnTyValys 60
Db 2495 ---ATCGGCTGTTCCCTTCTTAATCATGTGAG 2524

RESULT 15
US-09-023-655-1356
; Sequence 1356, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-D001 US

```

```

US-09-720-934-2_COPY_999_1062 (1-64) x US-09-023-655-1356 (1-1414)

QY      5 AlAGlnValIleAlaseerYrThraIathrglyProGluGlnLeuThrLeuAlaProGly 24
          ||| ||| |::||| ::||| |::||| |::||| |::||| |::||| |::||| |::|||
Db      432 GCTTATGTGAATTAACTTAACACTCAATGGCTGAGAAGAGATGAATTATCATTTAATAAGGGG 491
QY      25 GluLeuIleLeuIleArGlysLysAsnProGlyGlyTPPTPGluGlyGluLeuGlnAla 44
          ::::~::~: ~::||| |::||| |::||| |::||| |::||| |::||| |::|||
Db      492 ACAAGGTGATCGTCATGAGAAATGCACAAGTAGTGATGGTGGCCGTGGCACTAACATGA 551
QY      45 ArgGlyLysLysArgGlnIleGlyTrpPheProAlaAsnTYrVal 59
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      552 -----CAAAGTGGATGGTCCCTTCAAACTATGA 581

Alignment Scores:
Pred. No.:           6.04e-05              Length:         1414
Score:             107.50                  Matches:          20
Percent Similarity: 54.55%                 Conservative:    10
Best Local Similarity: 36.36%               Mismatches:     20
Query Match:       31.52%                   Indels:          5
Db:                4                       Gaps:            1

TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ. ID NO: 1356:
SEQUENCE CHARACTERISTICS:
LENGTH: 1414 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g35014
US-09-023-655-1356

Search completed: July 1, 2004, 19:53:56
Job time : 49.815 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:09:20 ; Search time 196.347 Seconds
(without alignments)
1572.471 Million cell updates/sec

Title: US-09-720-934-2_COPY_999_1062
Perfect score: 341
Sequence: 1 GEEIAQVIASVATGPEQLT.....RKKKRQIGFNPANYKLSP 64

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3163042 seqs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool.p/US09720934/runat.30062004.064541.13481/app.query.fasta_1.1386
-DB=Published Applications NA -GEMT=fastap -SUFFIX=rmpb -MISMATCH=0.1
-DOOPCL=0 -DOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USRR=US09720934 @CGN 1 1500 @runat.30062004.064541.13481
-NCPU=6 -ICPU=3 -NO MMAP -LARGEOQUERY -NEG SCORES=0 -WAIT -DSPBIOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-RCAPOP=6 -RCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	319	93.5	2067	16	US-10-264-049-887	Sequence 887, App
2	253	74.2	747	9	US-09-879-957-39	Sequence 39, Appl
3	253	74.2	2873	9	US-09-879-957-193	Sequence 193, App
4	253	74.2	3746	11	US-09-764-875-176	Sequence 176, App
5	253	74.2	4053	13	US-10-342-887-1882	Sequence 1882, Ap
6	253	74.2	4053	13	US-10-172-118-1882	Sequence 1882, Ap
7	253	74.2	4210	9	US-09-764-868-125	Sequence 125, App
8	253	74.2	5828	13	US-10-398-885A-15	Sequence 15, Appl
9	117.5	34.5	852	15	US-10-259-165-471	Sequence 471, App
10	117.5	34.5	855	15	US-10-259-165-143	Sequence 143, App
11	117.5	34.5	1563	17	US-10-437-963-41013	Sequence 41013, A
12	116.5	34.2	564	15	US-10-029-386-15	Sequence 15, Appl
13	114.5	33.6	361	13	US-10-425-114-5871	Sequence 5871, Ap
14	114.5	33.6	1420	13	US-10-425-114-5871	Sequence 31465, A
15	114.5	33.6	1505	13	US-10-425-114-25238	Sequence 25238, A
16	114.5	33.6	1514	13	US-10-425-114-28939	Sequence 28939, A
17	114.5	33.6	1559	13	US-10-425-114-25722	Sequence 25722, A
18	113.5	33.3	386	13	US-10-085-783A-37196	Sequence 37196, A
19	113.5	33.3	386	16	US-10-242-535A-37196	Sequence 37196, A
20	113.5	33.3	400	13	US-10-085-783A-22639	Sequence 22639, A
21	113.5	33.3	400	16	US-10-242-535A-22639	Sequence 22639, A
22	113.5	33.3	1215	17	US-10-648-593-61	Sequence 61, Appl
23	113.5	33.3	3143	14	US-10-144-621-1	Sequence 1, Appl
24	113.5	33.3	3348	13	US-10-342-887-187	Sequence 187, App
25	113.5	33.3	3348	13	US-10-172-118-187	Sequence 187, App
26	112.5	33.0	1582	15	US-10-259-165-580	Sequence 580, App
27	112.5	33.0	1661	13	US-10-425-114-34107	Sequence 34107, A
28	111	32.6	518	9	US-09-796-692-5870	Sequence 5870, Ap
29	111	32.6	518	15	US-10-040-862-5870	Sequence 5870, Ap
30	111	32.6	518	16	US-10-057-475B-5870	Sequence 5870, Ap
31	111	32.6	518	16	US-10-154-884B-5870	Sequence 5870, Ap
32	111	32.6	2757	10	US-09-960-706-1059	Sequence 1059, Ap
33	109.5	32.1	601	15	US-10-259-165-494	Sequence 494, App
34	107.5	31.5	1380	13	US-10-087-192-140	Sequence 140, App
35	107.5	31.5	1414	13	US-10-342-887-1250	Sequence 1250, Ap
36	107.5	31.5	1414	13	US-10-172-118-1250	Sequence 1250, Ap
37	107.5	31.5	1414	17	US-10-641-643-1356	Sequence 1356, Ap
38	107.5	31.5	1513	15	US-10-210-120-20	Sequence 20, Appl
39	107.5	31.5	3227	16	US-10-115-831-10	Sequence 10, Appl
40	107.5	31.5	79831	13	US-10-087-192-139	Sequence 139, App
41	106.5	31.2	909	13	US-10-425-114-9122	Sequence 9122, Ap
42	106.5	31.2	971	9	US-09-879-957-197	Sequence 197, App
43	106.5	31.2	1534	13	US-10-087-192-143	Sequence 143, App
44	106.5	31.2	1603	13	US-10-424-599-11257	Sequence 131257, Sequence 123461,
45	106.5	31.2	1890	13	US-10-424-599-123461	

ALIGNMENTS

RESULT 1
US-10-264-049-887
; Sequence 887, Application US/10264049
; Publication No. US2004000579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264, 049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 887
; LENGTH: 2067
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5)..(5)

OTHER INFORMATION: n equals a,t,b, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (2058) ..(2058)
OTHER INFORMATION: n equals a,t,b, or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (2063) ..(2063)
OTHER INFORMATION: n equals a,t,b, or c
US-10-264-049-887

Alignment Scores:
Pred. No.: 3.35e-35
Score: 319.00
Percent Similarity: 98.39%
Best Local Similarity: 95.16%
Query Match: 93.55%
DB: 16

Length: 2067
Matches: 59
Conservative: 2
Mismatch: 1
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x US-10-264-049-887 (1-2067)

QY 3 Glu1leAlaGluVal1leAlaSerTyrThrAlaThrGlyProGluGlnLeuThrlleuAla 22
Db 62 AAATGCCCCAGGTATTGCTTATACACCGCCCGCCGAGCAGCTCCTCGCC 121

QY 23 ProGlyGlnLeu1leAlaSerTyrThrAlaThrGlyProGluGlnLeuThrlleuAla 42
Db 122 CCGTGTACGCTGATTGATTCGAAAAAGAACCCAGGTGATGCTGGAGAGAGCTG 181

QY 43 Glu1leAlaGlyLysLysArgGln1leGlyTyrPheProAlaSerTyrValLysLeuLeu 62
Db 182 CAAGCAGGTGGAAAAAGCGCCAGATAGGCTGTTCCAGCTATATGTAAAGCTTCTA 241

QY 63 SerPro 64
Db 242 AGCCCT 247

RESULT 2

US-09-879-957-39
Sequence 39, Application US/09879957
Patent No. US20020034755A1
GENERAL INFORMATION:

APPLICANT: SPARKS, Andrew B.
HOFMAN, No. US20020034755A1h
KAY, Brian K.
FOWLES, Dana M.

TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
USING SAME

NUMBER OF SEQUENCES: 227

CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/879,957
FILING DATE: 13-Jun-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/630,915
FILING DATE: 03-APR-1996

ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 747 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-09-879-957-39

Alignment Scores:
Pred. No.: 2.32e-26
Score: 253.00
Percent Similarity: 87.10%
Best Local Similarity: 75.81%
Query Match: 74.19%
DB: 9
Length: 747
Matches: 47
Conservative: 7
Mismatch: 8
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x US-09-879-957-39 (1-747)

QY 3 Glu1leAlaGluVal1leAlaSerTyrThrAlaThrGlyProGluGlnLeuThrlleuAla 22
Db 334 GAGATTCTCAGTAACCTTCACTATGCTTCTGTTCTGAACAACCTAAGCTTGA 393

QY 23 ProGlyGlnLeu1leAlaSerTyrThrAlaThrGlyProGluGlnLeuThrlleuAla 42
Db 394 CCAGACAGTTAATTAATTCTAAGAAAAATCAAGTGGTGTGTCGCAAGAGAGTTA 453

QY 43 Glu1leAlaGlyLysLysArgGln1leGlyTyrPheProAlaSerTyrValLysLeuLeu 62
Db 454 CAGGCCAGAGAAAAAGCAGACAGAAAGATGTTCTCCAGCTCATGTTAACTTTTG 513

QY 63 SerPro 64
Db 514 GGTCCA 519

RESULT 3

US-09-879-957-193
Sequence 193, Application US/09879957
Patent No. US20020034755A1
GENERAL INFORMATION:

APPLICANT: SPARKS, Andrew B.
HOFMAN, No. US20020034755A1h
KAY, Brian K.
FOWLES, Dana M.

TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
USING SAME

NUMBER OF SEQUENCES: 227

CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/879,957
FILING DATE: 13-Jun-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/630,915


```

; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 41013
; LENGTH: 1563
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_4439C.1
US-10-437-963-41013

Alignment Scores:
Pred. No.: 9,12e-07
Score: 117.50
Percent Similarity: 59.65%
Best Local Similarity: 43.86%
Query Match: 34.46%
DB: 17
Length: 1563
Matches: 25
Conservative: 9
Mismatch: 18
Indels: 5
Gaps: 2

US-09-720-934-2_COPY_999_1062 (1-64) x US-10-437-963-41013 (1-1563)
QY 4 1LEAAGInVal11lea1aSerTYrThra1aThrglyProgluInleuThrLeuAlaPro 23
DB 1093 TTGGCTAGGCGCACTCCAGTCTTATCGTGGAGAGTGAAMACTGCTCAACCTGGCAGCT 1152
QY 24 G1Glnleu1leu1leu1lea1aArg1yLysAsnProglYg1yTTPg1uG1uLeuGln 43
DB 1153 GGTCACATATATAGTGTGCGGACAGCTGTCAACATGATGGAGGAGGTGAA----- 1206
QY 44 A1aArg1yLys1yLysAArg1n1leG1yTTPhePro1aAsnTYrVal1ys 60
DB 1207 TGCAGAGGGGAAA-----GCTGGCTGGTTCCTTACGACTACATCGAG 1248

RESULT 12
US-10-029-386-15/C
; Sequence 15, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David K.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEWOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL135783.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.55
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.52
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.51
; OTHER INFORMATION: SWISSPROT HIT Q14155, EVALUE 6.00e-27
; OTHER INFORMATION: NT HIT: g14745025, EVALUE 1.00e-110
; OTHER INFORMATION: EST_HUMAN HIT F20693.1, EVALUE 0.00e+00
US-10-029-386-15

Alignment Scores:
Pred. No.: 3.82e-07
Score: 116.50
Percent Similarity: 58.82%
Length: 564
Matches: 19
Conservative: 11
```

```

Best Local Similarity: 37.25%
Query Match: 34.16%
DB: 15
Indels: 5
Gaps: 1

US-09-720-934-2_COPY_999_1062 (1-64) x US-10-029-386-15 (1-564)
QY 10 SerTYrThra1aThrglyProgluInleuThrLeuAlaProglYg1nleu1leu1le 29
DB 408 AACTTAAAGCAGACTAATGAGATGAACGTGCTGAGTGTGAAGGAGGACATCAATTAAGCTC 349
QY 30 Arg1yLysAsnProglYg1yTTPg1uG1uLeuGln1aArg1yLys1yArg 49
DB 348 ACAGAGTGTGAAGAGAGAGGCTGTGGGAGGACATTAAATGGGAG----- 301
QY 50 G1n1leG1yTTPhePro1aAsnTYrVal1ys 60
DB 300 ---ACAGCTGTGTTCCCGAGTAATATATGTCCT 271

RESULT 13
US-10-425-114-5871
; Sequence 5871, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53113)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 5871
; LENGTH: 361
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700476119_FLI
US-10-425-114-5871

Alignment Scores:
Pred. No.: 4.35e-07
Score: 114.50
Percent Similarity: 59.65%
Best Local Similarity: 42.11%
Query Match: 33.58%
DB: 13
Length: 361
Matches: 24
Conservative: 10
Mismatch: 18
Indels: 5
Gaps: 2

US-09-720-934-2_COPY_999_1062 (1-64) x US-10-425-114-5871 (1-361)
QY 4 1LEAAGInVal11lea1aSerTYrThra1aThrglyProgluInleuThrLeuAlaPro 23
DB 36 TTGGCTAGGCGCACTCCAGTCTATACCAAGCCGAAAGCATACCGAGCTCAACCTTCAACT 95
QY 24 G1Glnleu1leu1leu1lea1aArg1yLysAsnProglYg1yTTPg1uG1uLeuGln 43
DB 96 GGTCACATATATAGTGTGCGGACAGTGTCAACATGATGGCCGAGGTGAA----- 149
QY 44 A1aArg1yLys1yLysAArg1n1leG1yTTPhePro1aAsnTYrVal1ys 60
DB 150 TGCAGAGGGGAAA-----GCTGGTGGTTCCTTACGACTACATCGAG 191

RESULT 14
US-10-425-114-31465
; Sequence 31465, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
```

```
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ. ID NOS: 73128
/ SEQ. ID NO 31465
/ LENGTH: 1420
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: UC-ZMFLB73177F01_FLI
US-10-425-114-31465

Alignment Scores:
Pred. No.: 2,17e-06 length: 1420
Score: 114.50 Matches: 22
Percent Similarity: 57.89% Conservative: 11
Best Local Similarity: 38.60% Mismatches: 19
Query Match: 33.58% Indels: 5
DB: 13 Gaps: 1

US-09-720-934-2_COPY_999_1062 (1-64) x US-10-425-114-31465 (1-1420)
QY 4 11eAlGlnVal11eAlaSerTyrThrAlaThrGlyProGlnLeuThrLeuAlaPro 23
Db 924 CTGGCTGAGGCAATCCATCATATCGAGCAGACGACACGTGAGCTCAACCTTTCAGCT 983
QY 24 G1yGlnLeu11eLeu11eArgLysLysAsnProG1yG1yTrrpGlnG1yGlnLeuGln 43
Db 984 GGTGACTACATTAAGTGTCCAAAGGTGTGCAACAAATGATGGCCGAAAGGTGAATGCAAG 1043
QY 44 AlaArgG1yLysLysArgGln11eG1yTrrpPheProAlaAsnTyrValLys 60
Db 1044 GGGAGAGCT-----GGCTGGTTCCTCAGACTACATCGAG 1079

RESULT 15
US-10-425-114-25238
/ Sequence 25238, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ. ID NOS: 73128
/ SEQ. ID NO 25238
/ LENGTH: 1505
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB732-023-C2_FLI
US-10-425-114-25238

Alignment Scores:
Pred. No.: 2.32e-06 length: 1505
Score: 114.50 Matches: 24
Percent Similarity: 59.65% Conservative: 10
Best Local Similarity: 42.11% Mismatches: 18
Query Match: 33.58% Indels: 5
DB: 13 Gaps: 2

US-09-720-934-2_COPY_999_1062 (1-64) x US-10-425-114-25238 (1-1505)
```

```
QY 4 11eAlGlnVal11eAlaSerTyrThrAlaThrGlyProGlnLeuThrLeuAlaPro 23
Db 1115 TTGGCTGAGGCAATCCATCATATCCAAAGCGAATACCGAGCTCAACCTTTCAGCT 1174
QY 24 G1yGlnLeu11eLeu11eArgLysLysAsnProG1yG1yTrrpGlnG1yGlnLeuGln 43
Db 1175 GGTGACTACATTAAGTGTCCAAAGGTGTGCAACAAATGATGGCCGAAAGGTGA 1228
QY 44 AlaArgG1yLysLysArgGln11eG1yTrrpPheProAlaAsnTyrValLys 60
Db 1229 TGCAGGGGGAAA-----GCTGGTGGTTCCTCAGACTACATCGAG 1270
```

Search completed: July 1, 2004, 20:16:08
Job time : 202.347 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:01:26 ; Search time 1925.25 Seconds

(without alignments)
992.694 Million cell updates/sec

Title: US-09-720-934-2_COPY_999_1062
Perfect score: 341
Sequence: 1 GEEIQVASTATATGREQUT.....RQKKQIGFNPANYKLLSP 64

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool.p/US09720334/runat 30062004 064540 13442/app query.fasta_1.1386
-DB=EST -QFMT=fastap -SUFFIX=est -MIMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=-1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45
-DOCCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAVISIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09720334.@CGN 1 1 12421 @runat 30062004 064540 13442 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em_escba:*
2: em_esthum:*
3: em_escin:*
4: em_estmu:*
5: em_estov:*
6: em_escpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_esc3:*
13: gb_esc4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_hiv:*
19: em_gss_pin:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pig:*
27: em_gss_vr1:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	330	96.8	538	14	CA390605	cs11id07
2	330	96.8	545	14	CB161912	CB161912 K-EST0221
3	330	96.8	583	14	CB156803	CB156803 K-EST0215
4	330	96.8	600	9	AL134506	AL134506 DKFZP547M
5	330	96.8	628	13	BQ393010	BQ393010 NISC mq28
6	330	96.8	629	14	CF768246	CF768246 CBS003921
7	330	96.8	750	14	CD356662	CD356662 AGENCOURT
8	330	96.8	1062	13	BQ898718	BQ898718 AGENCOURT
9	325	95.3	446	29	CG508946	CG508946 CEST59780
10	325	95.3	537	9	AA063751	AA063751 m379g10.y
11	325	95.3	603	9	AI594919	AI594919 UI-M-EVO-
12	325	95.3	610	13	BQ443095	BQ443095 UI-M-EVO-
13	325	95.3	694	13	BQ054334	BQ054334 UI-M-EVO-
14	325	95.3	732	12	BM950476	BM950476 UI-M-EVO-
15	325	95.3	782	13	BQ614534	BQ614534 UI-M-EVO-
16	325	95.3	875	10	BF122946	BF122946 601761612
17	325	95.3	898	14	CA987804	CA987804 AGENCOURT
18	325	95.3	906	13	BQ519029	BQ519029 AGENCOURT
19	325	95.3	2589	11	AK082606	AK082606 Mus muscu
20	325	95.3	5385	11	BC062938	BC062938 Mus muscu
21	318	93.3	575	14	CB163763	CB163763 K-EST0224
22	317	93.0	636	14	CD347867	CD347867 UI-M-EVO-
23	317	93.0	633	14	CA353372	CA353372 624874 NC
24	317	93.0	760	13	BX879964	BX879964 BX879964
25	317	93.0	805	13	BQ451981	BQ451981 603767625
26	311	91.2	668	14	CF763755	CF763755 CBS005445
27	299	87.7	346	14	R39287	R39287 ytc69g01.t1
28	286	83.9	522	28	AQ252614	AQ252614 HS 2010.B
29	268	78.6	298	14	F06100	F06100 HSC0XA031 n
30	265	77.7	518	28	A07229475	A07229475 HS 2154.B
31	266	75.1	590	9	AV866072	AV866072 AV866072
32	266	75.1	637	13	BM232778	BM232778 BM232778
33	266	75.1	709	9	AV849639	AV849639 AV849639
34	256	75.1	712	13	BM092285	BM092285 BM092285
35	253	74.2	362	12	BG089231	BG089231 uc64d07.y
36	253	74.2	463	13	BX504944	BX504944 DKFZP686G
37	253	74.2	1503	11	AK054527	AK054527 Mus muscu
38	252	73.9	785	13	BX276672	BX276672 BX276672
39	249	73.0	652	14	CF766841	CF766841 CBS003922
40	248	72.7	274	14	F07724	F07724 HSC2FG081 n
41	244	71.6	443	13	BQ430865	BQ430865 UI-HR-BNO
42	243	71.3	764	13	BU106683	BU106683 602554522
43	237	69.5	418	14	CF144063	CF144063 UI-HR-BPO
44	236	69.2	309	9	AJ281124	AJ281124 4A3A-ABB-
45	236	69.2	408	9	AJ283883	AJ283883 4A3B-AAG-

ALIGNMENTS

RESULT 1
LOCUS CA390605 538 bp mRNA linear EST 06-NOV-2002
DEFINITION cs11id07.y1 Human Retinal pigment epithelium/choroid cDNA
(Un-normalized, unamplified): cs Homo sapiens cDNA clone cs11id07
5', mRNA sequence.
ACCESSION CA390605
VERSION CA390605.1 GI:24721774
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 538)
 AUTHORS Wistow,G., Bernsteins,S.L., Wyatt,M.K., Farris,R.N., Behal,A.,
 Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
 TITLE Expressed sequence tag analysis of human RPE/choroid for the
 NEIBank Project: Over 6000 non-redundant transcripts, novel genes
 and splice variants
 JOURNAL Mol. Vis. 8 (4), 205-220 (2002)
 MEDLINE 22103460
 PUBMED 12107410
 COMMENT Contact: Wistow G
 Section on Molecular Structure and Function
 National Eye Institute
 6/331, NIH, Bethesda, MD 20892-2740, USA
 Tel: 301 402 3452
 Fax: 301 496 0078
 Email: graeme@helix.nih.gov
 Plate: 111 row: d column: 07
 Seq primer: M13RPI reverse primer (ABI).
 FEATURES
 SOURCE
 1. 538
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="cs111d07"
 /tissue_type="RPE/choroid"
 /dev_stage="Adult"
 /lab_host="EMD10B"
 /clone_lib="Human Retinal pigment epithelium/choroid cDNA
 (un-normalized, unamplified): cs"
 /note="Organ: Eye; Vector: pCMVSPORT6; Two different donor
 eyes (75-80 years old) yielded approximately 600 mg of
 dissected RPE/choroid tissue. This in turn yielded 340 ug
 of total RNA and 7 ug of mRNA. A directionally cloned cDNA
 library in the pCMVSPORT6 vector was constructed at Life
 Technologies (Rockville, MD; now part of Invitrogen Corp),
 essentially following the protocols of the SuperScript
 Plasmid System (Invitrogen Corp).
 <http://www.invitrogen.com/>. The library code
 designation was cs. For this library, cDNA inserts were
 cloned into the NotI/MluI sites of the vector. EST
 analysis was performed on the unamplified library at the
 NIH Intramural Sequencing Center (NISC)."

ORIGIN
 Alignment Scores:
 Pred. No.: 2.27e-28 Length: 538
 Score: 330.00 Matches: 62
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.77% Indels: 0
 DB: 14 Gaps: 0
 US-09-720-934-2_COPY_999_1062 (1-64) x CM390605 (1-538)

QY 3 GlnlleaaglnvallealeaserTyrThralaThrglyProgluglnleuThrlenua 22
 Db 131 GAAATTCGCCAGGTATTGCTCATACACCGCCACCGCCCGAGCAGCTACTCTGCC 190
 QY 23 Proglglnleuileleuileuileuileuileuileuileuileuileuileu 42
 Db 191 CTTGTGACGTGATTGATTCGAAAAAGAACCCAGGTGATGTGGGAAGAGAGCTG 250
 QY 43 GlnlaaarglyslvsarglnlleglYTrpPhepolaasnyYrvallysleu 62
 Db 251 CAAACAGCTGGGAAAAAGCGCCAGATAGGCTGTTCCAGCTAATATGTAAGCTTCTA 310
 QY 63 SerPro 64
 Db 311 AGCCCT 316
 RESULT 2
 CBI61912
 LOCUS CBI61912 545 bp mRNA linear EST 30-JAN-2003

DEFINITION K-EST0221969 L17N670205n1 Homo sapiens cDNA clone
 L17N670205n1-41-H01 5', mRNA sequence.
 ACCESSION CBI61912
 VERSION CBI61912.1 GI:28148038
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 545)
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoseun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 41 row: H Column: 01
 High quality sequence stop: 545.
 FEATURES
 SOURCE
 1. 545
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="L17N670205n1-41-H01"
 /sex="F"
 /lab_host="TOP10F"
 /clone_lib="L17N670205n1"
 /note="Organ: Liver; Vector: pT73-Pac; Site_1: EcORI;
 Site_2: NotI; The library was contributed by the Soares
 laboratory and it was constructed as described by Bonaldo,
 M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
 6(9): 791-806. RNA was prepared from harvested cell
 culture."

ORIGIN
 Alignment Scores:
 Pred. No.: 2.3e-28 Length: 545
 Score: 330.00 Matches: 62
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.77% Indels: 0
 DB: 14 Gaps: 0
 US-09-720-934-2_COPY_999_1062 (1-64) x CBI61912 (1-545)

QY 3 GlnlleaaglnvallealeaserTyrThralaThrglyProgluglnleuThrlenua 22
 Db 124 GAAATTCGCCAGGTATTGCTCATACACCGCCACCGCCCGAGCAGCTACTCTGCC 183
 QY 23 Proglglnleuileleuileuileuileuileuileuileuileuileuileu 42
 Db 184 CTTGTGACGTGATTGATTCGAAAAAGAACCCAGGTGATGTGGGAAGAGAGCTG 243
 QY 43 GlnlaaarglyslvsarglnlleglYTrpPhepolaasnyYrvallysleu 62
 Db 244 CAAACAGCTGGGAAAAAGCGCCAGATAGGCTGTTCCAGCTAATATGTAAGCTTCTA 303
 QY 63 SerPro 64
 Db 304 AGCCCT 309
 RESULT 3
 CBI56803
 LOCUS CBI56803 583 bp mRNA linear EST 29-JAN-2003
 DEFINITION K-EST0215771 L17N670205n1 Homo sapiens cDNA clone
 L17N670205n1-4-C07 5', mRNA sequence.
 ACCESSION CBI56803

VERSION	CB156803.1	GI:28141931
KEYWORDS	EST	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Bukacinska; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 583)	
AUTHORS	Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.U., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.	
TITLE	21C Frontier Korean EST Project 2001	
JOURNAL	Unpublished (2002)	
COMMENT	Contact: Kim YS Genome Research Center Korea Research Institute of Bioscience & Biotechnology 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea Tel: +82-42-860-4470 Fax: +82-42-860-4409 Email: yongsung@mail.kribb.re.kr Plate: 4 row: C column: 07 High quality sequence stop: 583.	
FEATURES	Location/Qualifiers	
source	1..583 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="L17N670205n1-4-C07" /sex="F" /lab_host="T0p10F" /clone_idb="L17N670205n1" /note="Organ: Liver; Vector: pT7T3-Pac; Site 1: EcoRI; Site 2: NotI; The library was contributed by the Soares laboratory and it was constructed as described by Donald, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cell culture."	
ORIGIN		
Alignment Scores:		
Prod. No.:	2.5e-28	Length: 583
Score:	330.00	Matches: 62
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	96.77%	Indels: 0
DB:	14	Gaps: 0
US-09-720-934-2_COPY_999_1062 (1-64) x CB156803 (1-583)		
OY	3	GluUlealagInValIleIleISerYrThralaThrGlyProGluGlnLeuThrIleuA 22
Db	124	GAATTCGCCCGAGTATTGCTCATACACCGCCACCGCCCGAGCAGCTCAGCTCTGCC 183
OY	23	ProGlyGlnLeuIleLeuIleIleArgYrYsYsAsnProGlyGlyYrPTPrGluGlyGlu 42
Db	184	CCTGGTCAGCTGATTTTATCCGAAAAAGAACCCAGGTGATGTGGAAAGAGAGCTG 243
OY	43	GlnAlaArgGlyLysYsArgGlnIleGlyTrpPheProAlaAsnYrValYsLeu 62
Db	244	CAGGAGCTGGGAAAAAGCCGACAGTACGCTGTTCCAGCTAATTATGTAAGCTCTTA 303
OY	63	SerPro 64
Db	304	AGCCCT 309
RESULT 4		
LOCUS	AL134506	600 bp mRNA linear EST 04-SEP-2003
DEFINITION	DKFZP547M065_r1 547 (synonym: hfpri1) Homo sapiens cDNA clone	
ACCESSION	AL134506	DKFZP547M065_5', mRNA sequence.
KEYWORDS	AL134506.1	GI:6602693
VERSION	EST	
SOURCE	Homo sapiens (human)	

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 600)
AUTHORS Blum H., Bauesachs S., Mewes W., Weil B. and Wiemann S.
TITLE EST (Blum,H., Bauesachs,S., Mewes,H.W., Weil,B. and Wiemann,S.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS Infoltaedter Landstr.1, D-95764 Neuherberg, Germany

This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de,
sequenced by LMU (Ludwig Maximilians University,
Munich/Germany) within the cDNA sequencing consortium of the German
Genome Project.

No sl sequence available.
This clone (DKFZp547M065) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcententrum, Neuherweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..600
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp547M065"
/tissue_type="brain"
/dev_stage="fetal"
/lab_host="X1-2blue"
/clone_id="547 (synonym: hfbn1)"
/note="Vector: pAMP1; Site_1: NotI; Site_2: SalI"

ORIGIN

Alignment Scores:

Pred. No.: 2.58e-28 Length: 600
Score: 330.00 Matches: 62
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.77% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x AL134506 (1-600)

OY 3 GlulleaaglnValllleaasertYrThraAlathrglyProglugInleuTrlleuA 22
| | | | |
Db 260 GAATTCGCCAGGTATTGCTCATTAACCCGCCCGCCCCGAGCAGCTCACCTCGCC 319
| | | | |
OY 23 ProGlgnleuileuileuilearqLYsaAsnProGlygLyTPPTPnglugInleu 42
| | | | |
Db 320 CCTGGTAGCGTGAATTTGCATCCGAAAAAGAACCCAGGTGGATGTGGGAAGGAGACTG 379
| | | | |
OY 43 GlmlalaarglglylsysargGlnllcglyTTPheProhlaantYrVallvIsleu 62
| | | | |
Db 380 CAAGCAAGTGAGAAAAGCCGCAATAGGCTGTTCCTCAGCTAAATTATGAAAACTTCTA 439
| | | | |

OY 63 SerPro 64
| | | | |
Db 440 AGCCCT 445

RESULT 5
BQ393010 628 bp mRNA linear EST 22-MAY-2002
LOCUS BQ393010
DEFINITION NISC mg28c06.y1 NICHD XGC Emb5 Silurana tropicalis cDNA clone
IMAGE J5384843 5', mRNA sequence.
ACCESSION BQ393010
VERSION BQ393010.1 GI:21080687
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Silurana tropicalis
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 628)

AUTHORS NIH-XCG <http://image.llnl.gov/image/html/xenopuslib.info.shtml>.
TITLE National Institute of Child Health and Human Development, National Cancer Institute, Xenopus Gene Collection
JOURNAL Unpublished (2002)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 cDNA Library Preparation: cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC)
 Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Info@image.llnl.gov
 Plate: LHAM1980 row: F column: 12
 Seq primer: M13RPI reverse primer (ABI).
FEATURES Location/Qualifiers
 source
 1..628
 /organism="Silurana tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="IMAGE:584843"
 /tissue_type="gastrula"
 /dev_stage="embryo, stages 10-13"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NICHD XCG Emb5"
 /note="Vector: PCMV-SPOB6.1; Site 1: NotI; Site 2: EcoRV; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.0 Kb. Constructed by Invitrogen. Note: This is a Xenopus Gene Collection (XGC) library."
ORIGIN
 Alignment Scores:
 Pred. No.: 2.73e-28 Length: 628
 Score: 330.00 Matches: 62
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.77% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_999_1062 (1-64) x BQ393010 (1-628)
 QY 3 GtlllealaglnvalillealaserTyrThrAlaThrglyProgluglnleuThrlenuA 22
 Db 416 GAATTCGCCAGGTATTGCTTTACACAGCAACTGGCCTGAGCAGCTGACACTGGCT 475
 QY 23 Proglglnleuileuileuileuileuileuileuileuileuileuileuileu 42
 Db 476 CTTGGCAACTTATCTCATCTCGAAAAAGATCTCGAGAGATGCTGGAGAGAACTG 535
 QY 43 GlnAlaArgGlyLysLysArgGlnlleGlyTrrpPheProAlaAsnTrrValyleu 62
 Db 536 CAGCGAGCTGGAGAAAGCGACAGATTGATGTTTCCAGCAATATGTAAGCTTTTG 595
 QY 63 SerPro 64
 Db 596 AGCCCT 601
RESULT 6
 CF768246 629 bp mRNA linear EST 17-OCT-2003
 LOCUS CFS003921 Bos taurus skin cDNA library Bos taurus cDNA clone
 DEFINITION CCL003921 5', mRNA sequence.
 ACCESSION CF768246
 VERSION CF768246 GI:37717465
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 629)
AUTHORS Wang, Y.H., McWilliam, S., and Lehnert, S.
TITLE Transcription profiling of cattle skin

JOURNAL Unpublished (2003)
COMMENT Contact: Dr Yonghong Wang
 Functional Genomics Lab
 CSIRO Livestock Industries
 Level 5, Queensland Biosciences Precinct,
 306 Carmody Road St. Lucia QLD Australia
 Tel: 07 3214 2445
 Fax: 07 3214 2685
 Email: Yonghong.Wang@csiro.au
 Plate: 5 row: A column: 11.
FEATURES Location/Qualifiers
 source
 1..629
 /organism="Bos taurus"
 /mol_type="mRNA"
 /strain="Hereford Shorthorn"
 /db_xref="taxon:9913"
 /clone="CCL003921"
 /sex="female"
 /tissue_type="pooled"
 /dev_stage="Adult"
 /lab_host="XLI-BlueMR strain"
 /clone_lib="Bos taurus skin cDNA library"
 /note="Organ: skin; Vector: Uni-ZAPXR; Site 1: EcoRI; Site 2: Xho I; library made from pooled skin of adult female Hereford-Shorthorn."
ORIGIN
 Alignment Scores:
 Pred. No.: 2.73e-28 Length: 629
 Score: 330.00 Matches: 62
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.77% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_999_1062 (1-64) x CF768246 (1-629)
 QY 3 GtlllealaglnvalillealaserTyrThrAlaThrglyProgluglnleuThrlenuA 22
 Db 21 GAATTCGCCAGGTATTGCTTTACACAGCACTGGCTCTGACAGCTTACCTGGCC 80
 QY 23 Proglglnleuileuileuileuileuileuileuileuileuileuileuileu 42
 Db 81 CCGGTGAGCTGATCTGATCGGAAAAAGAACCAAGTGTGATGCGGAGAGAGCTA 140
 QY 43 GlnAlaArgGlyLysLysArgGlnlleGlyTrrpPheProAlaAsnTrrValyleu 62
 Db 141 CAGCAGCTGGGAAAAAGCGCAGATAGCGTGTTCCTGCTAATTCGTAAACTTTTA 200
 QY 63 SerPro 64
 Db 201 AGCCCT 206
RESULT 7
 CD356862 750 bp mRNA linear EST 29-MAY-2003
 LOCUS AGENCOURT 14253331 NIH MGC 187 Homo sapiens cDNA clone
 DEFINITION IMAGE:30404497 5', mRNA sequence.
 ACCESSION CD356862
 VERSION CD356862 GI:31128273
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 750)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: NDCM190 row: C column: 02
 High quality sequence stop: 519.
 Location/Qualifiers

FEATURES

Source

```
1..750
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30404497"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1b="NIH MGC 187"
/Note="Organ: Blood vessels - aorta, basilar and artery;
vector: pDNR-LIB; Site_1: Sfil (ggccatgagcc); Site_2:
Sfil (ggcgctcgcc); 5' and 3' adaptors were used in
cloning as follows: 5' adaptor sequence:
5'-CACGCCATTATGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4 kb
(range 0.5-4.0 kb). 14/15 colonies contained inserts by
PCR. This library was enriched for full-length clones and
was constructed by Clontech Laboratories (Palo Alto, CA).
Note: this is a NIH MGC Library."
```

ORIGIN

Alignment Scores:
 Pred. No.: 3,38e-28 Length: 750
 Score: 330.00 Matches: 62
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.77% Indels: 0
 DB: 14 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x CD356862 (1-750)

```
QY 3 GtlttlaagllVallllealaserYrrrralathgYlProgluglneutrrleuaA 22
DB 357 GAATTGCCCAAGTTATTCCTCATACCGCCAGCCGCCGAGCAGCTCCTCGCC 416
QY 23 ProglYglnleuileuileuilearYlYsAsnProglYgYlYrrPrpgluglYglnleu 42
DB 417 CCGGTCCAGCTGATTTTGATCCGAAAAAGAACCCAGGTGGATGTTGGGAGAGAGCTG 476
QY 43 GlnAlarGgYlYlYsYsArGlnllleglYrrPhPrAlaAsnYrValYlsleu 62
DB 477 CAAGCAGCTGGGAAAAAGCCGAGATAGGCTGGTTCCAGCTATTATGTAAGCTTCTA 536
QY 63 SerPro 64
DB 537 AGCCCT 542
```

RESULT 8
 LOCUS BQ898718 1062 bp mRNA linear EST 16-AUG-2002
 DEFINITION AGENCOUPT 8118759 lupski_dorsal_root_ganglion Homo sapiens cDNA
 ACCESSION BQ898718
 VERSION BQ898718.1 GI:22290732
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1062)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLM13562 row: n column: 16
 High quality sequence start: 116
 High quality sequence stop: 760.
 Location/Qualifiers

FEATURES

Source

```
1..1062
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6180303"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_1b="Lupski dorsal root ganglion"
/Note="Vector: pCMV-SPOK6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-GCTAGCTCTAGATCGGAGCGCGCCGCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
```

ORIGIN

Alignment Scores:
 Pred. No.: 5.12e-28 Length: 1062
 Score: 330.00 Matches: 62
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 96.77% Indels: 0
 DB: 13 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x BQ898718 (1-1062)

```
QY 3 GtlttlaagllVallllealaserYrrrralathgYlProgluglneutrrleuaA 22
DB 337 GAATTGCCCAAGTTATTCCTCATACCGCCAGCCGCCGAGCAGCTCCTCGCC 396
QY 23 ProglYglnleuileuileuilearYlYsAsnProglYgYlYrrPrpgluglYglnleu 42
DB 397 CCGGTCCAGCTGATTTTGATCCGAAAAAGAACCCAGGTGGATGTTGGGAGAGAGCTG 456
QY 43 GlnAlarGgYlYlYsYsArGlnllleglYrrPhPrAlaAsnYrValYlsleu 62
DB 457 CAAGCAGCTGGGAAAAAGCCGAGATAGGCTGGTTCCAGCTATTATGTAAGCTTCTA 516
QY 63 SerPro 64
DB 517 AGCCCT 522
```

RESULT 9
 LOCUS CG508946 446 bp DNA linear GSS 01-OCT-2003
 DEFINITION OST59780 Mus musculus 129Sv/Ev Mus musculus genomic clone OST59780,
 genomic survey sequence.
 ACCESSION CG508946
 VERSION CG508946.1 GI:37291689
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE	1 (bases 1 to 446)	
AUTHORS	Zambrowicz, B.P., Abuhin, A., Ramirez-Solis, R., Richier, L.J., Pigrog, U., Beltrandel-Rib, H., Buxton, E.C., Edwards, J., Finch, R.A., Fridde, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, M., Jaling, C., Key, B.W., Jr., Kipp, P., Kohlhauff, B., Ma, Z.-Q., Markesich, D., Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z., Sparks, M.J., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N., Zhu, O., Person, C. and Sands, A.T.	
TITLE	Wnt1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention	
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)	
COMMENT	Contact: Zambrowicz BP OmitBank Lexicon Genetics Incorporated 4000 Research Forest Drive, The Woodlands, TX 77381, USA Email: materials@lexgen.com Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al (Nature, 1998 Apr 9;392(6676):1608-11). Class: Gene Trap.	
FEATURES	Location/Qualifiers	
SOURCE	1..446	
	/organism="Mus musculus"	
	/mol_type="genomic DNA"	
	/strain="129Sv/Ev"	
	/db_xref="taxon:10090"	
	/clone="OST5978D"	
	/cell_type="embryonic stem cell"	
	/clone_11b="Mus musculus 129Sv/Ev"	
ORIGIN		
Alignment Scores:		
Pred. No.:	6,91e-28	Length: 446
Score:	325.00	Matches: 61
Percent Similarity:	98.39%	Conservative: 0
Best Local Similarity:	98.39%	Mismatches: 1
Query Match:	95.31%	Indels: 0
DB:	29	Gaps: 0
US-09-720-934-2_COPY_999_1062 (1-64) x CG508946 (1-446)		
QY	3 Glnlleaaglnvallealeaseryrthralathrgylprogluglnleuthreuala	22
Db	6 GAATTCGCCAGGTTATTGCTTCTACAGCTGTACTGTGCCGAACAACCTGCGCT	65
QY	23 Proglglnleulleleullealeaglylsysasnp-roglgyltprtptglugluleu	42
Db	66 CCTGGCGCGCTGATTCGATCCGCAAAAAGAACCCAGGTGATGTGGAGAGAACTGT	125
QY	43 GlnlaaArglyLysLysAArgGlnlleglylrPheProAlaAsnTyValylsleu	62
Db	126 CAACTCGAGGAGAAAAAGCGCCAGTAAAGGTGTTCCAGCAATTAATCTCAACTTCTA	185
QY	63 SerPro 64	
Db	186 AGCCCC 191	
RESULT 10		
LOCUS	AA063751	537 bp mRNA linear EST 24-SEP-1999
DEFINITION	U799g10.1 Soares mouse p3MWR19.5 Mus musculus cDNA clone IMAGE:483370 5' similar to WP:C14F5.5 CE01784 SEM-5: C. ELEGANS SEM-5: , mRNA sequence.	
ACCESSION	AA063751	
VERSION	AA063751.1 GI:1557790	
KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 537)	
AUTHORS	Meira, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, U., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,	

Theising,B., Mylie,T., Lennon,G., Soares,B., Wilson,R. and
 Materzow,R.
 The WashU-HM1 Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HM1 Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LML ; contact the
 IMAGE Consortium (info@image.lml.gov) for further information.
 MGI:293114
 Possible reversed clone: similarity on wrong strand
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 364.
 Location/Qualifiers
 1..537
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:482370"
 /dev_stage="19.5 dpc total fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares mouse p3NMF19.5"
 /note="Vector: pTV713D (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTTACCAATCTGACGTGGGAGCGCGCATTTTCTTTTCTTTT 3'] ,
 double-stranded cDNA was size selected, ligated to Eco RI
 adaptors (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pTV73 vector
 (Pharmacia) . Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M. Patricia Bonaldo. RNA was kindly provided by
 Dr. Minoru Ko (Wayne State University)."

ORIGIN
 Alignment Scores:
 Pred. No.: 8.63e-28 Length: 537
 Score: 325.00 Matches: 61
 Percent Similarity: 98.39% Conservative: 0
 Best local Similarity: 98.39% Mismatches: 1
 Query Match: 95.31% Indels: 0
 DB: Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x AA063751 (1-537)
 QY 3 GlnlelelagnivallealaserTyThrAlathrgelyProGInleuThrLeuAla 22
 Db 175 GAATTTCCCAAGGATTATGCTTCTTACGCGCTACTGATCCGACAACATCACCCTGGCT 234
 QY 23 ProGlyGInleuileleuileuilearxylslyAsnProSlyglyTyTPTPGInglyGInleu 42
 Db 235 CCTGGCAGCTGATTCATTCGATCCGAAAAAAGAACCAAGTGATGTGGGAAGGAACTGTG 294
 QY 43 GlAlAlarGglylyslsArGgInleGlyTTrPheProAlaAsnTyrrVallyLeuLeu 62
 Db 295 CAAGCTCGAGGGAAGAAAGCCGCGATGAGGTGGTTCCTCAGCAAAATATATGTAACACTTCTA 354
 QY 63 SerPro 64
 Db 355 AGCCCC 360

RESULT 11
 LOCUS A1594919 603 bp mRNA linear EST 15-MAR-2000
 DEFINITION m179g10.y1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
 IMAGE:482370 5' similar to TR:042287 042287 INTERSECTION. ;, mRNA
 sequence.
 ACCESSION A1594919
 VERSION A1594919.1 GI:4603967

KEYWORDS	EST
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scurionathi; Muridae; Murinae; Mus. 1 (bases 1 to 603)
AUTHORS	Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE	The WashU-NCI Mouse EST Project 1999
JOURNAL	Unpublished (1999)
COMMENT	Contact: Marra M/WashU-NCI Mouse EST Project 1999 Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watsen.wustl.edu This clone is available royally-free through LNL, contact the IMAGE Consortium (info@image.lnl.gov) for further information. This read is a RESEQUENCE of a previously sequenced mouse clone. This read has been verified (found to hit its original self in the correct orientation) MGI:293114 Seq primer: -40RP from Gibco High quality sequence stop: 464 POLYA=No.
FEATURES	location/Qualifiers 1..603 /organism="Mus musculus" /mol_type="mRNA" /db_xref="taxon:10090" /clone="IMAGE:482370" /dev_strage="19.5 dpc total fetus" /lab_host="MD10B (ampicillin resistant)" /clone_lib="Scars mouse p3MMP19.5" /note="Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - 0190(dT) primer (5' TGGTACCATCTGAAAGTGGAGCGCGCCATTTTTTTTTTTTTTTT 3'), double-stranded cDNA was size selected, ligated to Eco RI adapters (pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Patima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."
ORIGIN	
Alignment Scores:	9.92e-28 Length: 603
Pred. No.:	325..00 Matches: 61
Percent Similarity:	98.39% Conservative: 0
Best local Similarity:	98.39% Mismatches: 1
Query Match:	95.31% Indels: 0
DB:	9 Gaps: 0
US-09-720-934-2_COPY_999_1062 (1-64) x A1594919 (1-603)	
QY	3 GtutleataglnvalillealaserlyrthrilaThrlygProgluglnleuThrlenuA 22
Db	175 GAANTGGCCAGGTATTGCTTCACCTGCTNATGTCGGAACAACCTACCTGCT 234
QY	23 ProglglnleuilleleuillearglyslsasnProglglylYTPTPgluglyGlnue 42
Db	235 CCTGGGCGCGATTCGATCCGGAAAAAGAACCCAGGTGATGCTGGGAAGGAACTG 294
QY	43 GlnlaaAGGlylYslYsArGlnlleglylTrpPheprolaasrlyValylsleuenu 62
Db	295 CAAGTCAGAGGAAAAAGCCCAAGTAAAGGAGGTTTCCAGCAAAATTATGTCAACTCTA 354
QY	63 Serpro 64

```

Db                355  AGCCCC 360  |||||
                                     |||||
RESULT 12
BQ443095          610 bp  mRNA  linear  EST 29-MAY-2002
LOCUS             BQ443095
DEFINITION        UI-M-EV0-bx-f-j-12-0-UI.r1 NIH BMAP_EV0 Mus musculus cDNA clone
VERSION           BQ443095.1  GI:21246207
KEYWORDS          IMAGE:5707499 5', mRNA sequence.
ACCESSION         BQ443095
SOURCE            Mus musculus (house mouse)
ORGANISM          Mus musculus
AUTHORS           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE             Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.
JOURNAL           NIH-MGC http://mgc.nci.nih.gov/.
COMMENT           1 (bases 1 to 610)
                  NIH-MGC http://mgc.nci.nih.gov/.
                  National Institutes of Health, Mammalian Gene Collection (MGC)
                  Unpublished (1999)
                  Contact: Robert Strausberg, Ph.D.
                  Email: sgabbs-remail.nih.gov
                  Tissue Procurement: Dr. James Lin, University of Iowa
                  cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
                  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
                  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
                  Clone Distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  This clone was contributed by the Brain Molecular Anatomy Project
                  (BMAP)
                  Seg primer: pyX-5.
                  Location/Qualifiers
                    1..610
                       /organism="Mus musculus"
                       /mol_type="mRNA"
                       /strain="C57BL/6"
                       /db_xref="taxon:10090"
                       /clone="IMAGE:5707499"
                       /tissue_type="whole brain"
                       /dev_stage="embryo 15.5 dpc"
                       /lab_host="DH10B (T1 phage resistant)"
                       /clone_lib="NIH_BMAP_EV0"
                       /note="Organ: brain; Vector: pyX-Asc, Site_1: EcoR I,
                  Site_2: Not I; The library was constructed according to
                  Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                  1996. Denatured mRNA was size fractionated on a 1% agarose
                  gel. First strand cDNA synthesis was primed with an
                  oligo-dT primer containing a Not I site. Double stranded
                  cDNA was size selected according to mRNA size fraction,
                  ligated with EcoR I adaptor, digested with Not I, and then
                  cloned directionally into pyX-Asc vector. The library tag
                  sequence located between the Not I site and the polyA
                  tail, is GTGCGTGGAA. This library was created for the
                  University of Iowa Mouse Brain Molecular Anatomy Project
                  (BMAP). 'Gene Discovery in the Developing Mouse Nervous
                  System', supported by National Institutes of Mental Health
                  (NIH), Hemin Chin, Ph.D., program coordinator."
ORIGIN
Alignment Scores:
Pred. No.:      1.01e-27      Length:      610
Score:          325.00        Matches:      61
Percent Similarity: 98.39%    Conservative: 0
Best Local Similarity: 98.39% Mismatches:      1
Query Match:    95.31%       Gaps:          13
DB:              13          Gaps:          0

US-09-720-934-2_COPY_999_1062 (1-64) x BQ443095 (1-610)
QY      3  GhuUeAagcVbAlIleAseTyrThAlahrgGyPogUgnleUThLeuAa 22
Db      310  GAAATGCGCAAGATTATTCCTCTAGCGCTACtGtGCTCCGGAACAACtACCCtGAGT 369

```

QY	23	ProclgylgmneullelellellaeglylsysksnProclgylytptprglwglyluleu	42
Db	370	CCTGGGACGTGATTTCATCCCGAAAAAGAACCCAGGTGATGTGTGGAAAGAGMACTG	429
QY	43	GlnlaaragcilylvlsyarsglnllgelyttrpnphepcalaasnyVallysleulu	62
Db	430	CAACTCGAGGGAAAAGCCGCAGATAGGGTGTTTCCAGCAAAATATNGTCAAACTTCTA	489
QY	63	SerPro 64 	
Db	490	AGCCCC 495	
RESULT 13			
BUD054334		694 bp	mRNA linear EST 26-AUG-2002
DEFINITION	UI-M-PD0-bzh-n-08-0-U. r1 NIH BMAP_FDO Mus musculus cDNA clone		
ACCESSION	IMAGE:6403783 5', mRNA sequence.		
VERSION	BUD054334		
KEYWORDS	BUD054334.1 GI:22494411		
SOURCE	EST.		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Mus musculus		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
JOURNAL	(bases 1 to 694)		
COMMENT	NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: rsraus@remail.nih.gov Tissue Procurement: Dr. James Linn, University of Iowa CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLU at: http://image.lnl.gov This clone was contributed by the Brain Molecular Anatomy Project (BMAP)		
FEATURES	Seq primer: PYX-5. Location/Qualifiers		
SOURCE	1..694 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6" /db_xref="taxon:10090" /clone="IMAGE:6403783" /tissue_type="whole brain" /dev_stage="embryo 12.5 dpc" /lab_host="DH10B (TI phage resistant)" /clone_lib="NIH BMAP_FDO" /note="Organ: brain; Vector: pYX-asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according to Bonadio, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-asc vector. The library tag sequence located between the Not I site and the polyA tail, is TGAGAGGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."		
ALIGNMENT SCORES:	1.17e-27	length: 694	
Pred. No.:	325.00	Matches: 61	
Percent Similarity:	98.39%	Conservative: 0	
Best Local Similarity:	98.39%	Mismatches: 1	

	Query Match:	95.31%	Indels:	0
DB:		13	Gaps:	0
US-09-720-934-2_COPY_999_1062 (1-64) x BU054334 (1-694)				
OY	3	Glu1leaaglnvllllealaserTyrThrAlathrglyProglInleuthLeuala	22	
Dd	120	GAATTCGCCCGGTATTATGCTCCTACGGCTCTACTGTTCCGAAACAATCAACCTGCCT	179	
OY	23	ProglInleuileulealearylsyasnproglygTYrPTTPGLUGlVgludeu	42	
Dd	180	CCTGGAGAGCTGATTCTTGATCCGAAAAAAGAACCCAGGTGATGTGGGAAGCAACTG	239	
OY	43	GlnAlaarglyLylysArgGlnlEGlYTrpPheproalaasnyrValylsleueu	62	
Dd	240	CAAGCTCGAGGAAAAGCCGCAATGAGGTGGTTCCAGCAAAATATGTCAAACCTCTA	299	
OY	63	SerPro 64		
Dd	300	AGCCCC 305		
RESULT 14				
LOCUS	BM950476	732 bp	mRNA	linear EST 14-MAR-2002
DEFINITION	UT-M-EHD0-hum-o-05-0-U1.r1 NIH_BMAP_EHD0 Mus musculus cDNA clone			
ACCESSION	BM950476			
VERSION	BM950476.1	GI:19434066		
KEYWORDS	EST.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 732) NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)			
REFERENCE	Contact: Robert Strausberg, Ph.D. Email: rgabbs@remail.nih.gov Tissue Procurement: Dr. James Lin, University of Iowa cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov			
AUTHORS	This clone was contributed by the Brain Molecular Anatomy Project (BMAP)			
TITLE	Seq primer: pyX-5.			
JOURNAL	Location/Qualifiers			
COMMENT	1..732			
	/organism="Mus musculus"			
	/mol_type="mRNA"			
	/strain="C57BL/6"			
	/db_xref="taxon:10090"			
	/clone="IMAGE:5687260"			
	/tissue_type="whole brain"			
	/dev_stage="embryo 18.5 dpc"			
	/lab_host="MDH10B (T1 phage resistant)"			
	/clone_lib="NIH BMAP_EHD0"			
	/note="Organ: Brain; Vector: pyX-Asc; Site_1: Ecor I; Site_2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with Not I, and then cloned directionally into pyX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is CAGCACGAC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous			

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 12:16:50 ; Search time 1345.81 Seconds
(without alignments)
1900.151 Million cell updates/sec

Title: US-09-720-934-2_COPY_1080_1138
Perfect score: 323
Sequence: 1 AYCQYGMVDTAQNDELIA.....GEVNGVGLFPBNYKLTLD 59

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 segs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q/cg2 1/USPTO_spool_p/US09720934/runat_30062004_064539_13430/app_query.fasta_1.1386
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-INITs-bits -START=1 -END=-1 -MATRIX=bls62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAVYSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720934.@CGN 1.1 7509 @runat_30062004_064539_13430 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_mu:*
21: em_ov:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB	ID	Description
1	323	100.0	1996	9	AF180522	AF180522 Homo sapi
2	323	100.0	3231	6	BD205037	BD205037 Isolated
3	323	100.0	3241	9	HS061166	U61166 Human SH3 d
4	323	100.0	4321	9	HS060584	BX58175 Homo sapi
5	323	100.0	5195	6	BD205035	BD205035 Isolated
6	323	100.0	5199	6	BD205033	BD205033 Isolated
7	323	100.0	5287	9	AF064243	AF064243 Homo sapi
8	323	100.0	5381	9	AF114488	AF114488 Homo sapi
9	323	100.0	5458	6	BD205034	BD205034 Isolated
10	323	100.0	6439	9	AF114487	AF114487 Homo sapi
11	323	100.0	7247	9	AF064244	AF064244 Homo sapi
12	323	100.0	12015	9	AP000312	AP000312 Homo sapi
13	323	100.0	100000	9	AP000193	AP000193 Homo sapi
14	323	100.0	114929	9	AP000050	AP000050 Homo sapi
15	323	100.0	151516	9	AP000117	AP000117 Homo sapi
16	323	100.0	171361	9	BS000184	BS000184 Pan trogl
17	323	100.0	184526	2	AC144367	AC144367 Papio anu
18	323	100.0	194157	2	AC144368	AC144368 Papio anu
19	323	100.0	340000	9	AF001718	AF001718 Homo sapi
20	322	99.7	78190	9	AC106760	AC106760 Homo sapi
21	322	99.7	179264	2	AC023112	AC023112 Homo sapi
22	319	98.8	203423	2	AC137544	AC137544 Gallus ga
23	317	98.1	142974	2	AC129075	AC129075 Felis cat
24	317	98.1	157107	2	AC140666	AC140666 Sus scrof
25	317	98.1	169774	2	AC138787	AC138787 Sus scrof
26	313	96.9	1133	10	AF169621	AF169621 Mus muscu
27	313	96.9	3723	10	AF132478	AF132478 Mus muscu
28	313	96.9	5145	10	AF132481	AF132481 Mus muscu
29	313	96.9	217131	10	AC126053	AC126053 Mus muscu
30	312	96.6	224721	2	AC139632	AC139632 Bos tauri
31	308	95.4	3812	10	AF132672	AF132672 Rattus no
32	308	95.4	4025	10	AF127798	AF127798 Rattus no
33	308	95.4	135924	2	AC115529	AC115529 Rattus no
34	307	95.0	178801	2	AC121075	AC121075 Canis fam
35	307	95.0	235426	2	AC121074	AC121074 Canis fam
36	296	91.6	76179	2	AL606725	AL606725 Dario rer
37	296	91.6	104334	5	AL606751	AL606751 Zebrafish
38	296	91.6	163197	5	BX005416	BX005416 Zebrafish
39	296	91.6	182740	2	AC138436	AC138436 Dario rer
40	296	91.6	225535	2	BX470235	BX470235 Dario rer
41	295	91.3	4103	5	AF032118	AF032118 Xenopus l
42	282	87.3	113146	2	AC138440	AC138440 Tetradon
43	282	87.3	127051	2	AC114895	AC114895 Tetradon
44	282	87.3	134408	2	AC139628	AC139628 Takifugu
45	282	87.3	136937	2	AC139627	AC139627 Takifugu

RESULT 1

ALIGNMENTS

AF180522 AF180522 1996 bp mRNA linear PRI 05-SEP-1999
LOCUS Homo sapiens intersectin short form 2 (ITSN) mRNA, partial cds.
DEFINITION AF180522
ACCESSION AF180522.1 GI:5823551
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 1996)
Tsuda, L.O., Kvasna, S.M., Skripkina, I.Y., Anoprienko, O.V.,
Slavov, D., Tassone, F., Rynditch, A.V. and Gardiner, K.
JOURNAL Mouse homologs of human chromosome 21 genes
TITLE Unpublished
REFERENCE 2 (bases 1 to 1996) Skripkina, I.Y., Anoprienko, O.V.,
Tsuda, L.O., Kvasna, S.M., Rynditch, A.V. and Gardiner, K.
AUTHORS Slavov, D., Tassone, F., Rynditch, A.V. and Gardiner, K.
JOURNAL Direct Submission
TITLE Submitted (25-AUG-1999) Department of Molecular Oncogenetics,
Institute of Molecular Biology and Genetics of National Academy of
Sciences of Ukraine, 150, Zabolotnogo str., Kiev 252627, Ukraine
FEATURES
Source Location/Qualifiers
1..1996
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
/tissue_type="brain"
/dev_stage="infant"
<1..1996
/gene="ITSN"
<1..566
/gene="ITSN"
/codon_start=3
/product="intersectin short form 2"
/protein_id="AAB53183.1"
/db_xref="GI:5823552"
/translation="KSYVRLISGPIRSTSMGSSSPASIKRVAAPKPVSGE
IAQYIASYTAAGPEQTLAPQGLIRKKNRGMWEGELQKRGKRGTFNPANYKL
LSPEKSKITPEPRKSTALAAVCQVIGMDYDTAQNDELAKKQIINVINKEDPDW
KGEVNGVGVPSPSNYKLTITMDPSQ"

Alignment Scores:
Pred. No.: 4.01e-32 Length: 1996
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x AF180522 (1-1996)

QY 1 AAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
Db 369 GCACTGTGCCAGCGATTGGAGTGAAGACTACACCGCGCAAGATGACGATGAGCTGGCC 428

QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTyrLysGly 40
Db 429 TTCAACAAGGGCCGATCATCAACGCTCTCAACAAGGAGGACCTGACTGCTGGAAAAGA 488

QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59
Db 489 GAAGTCATGACACAGTGGGCGCTTCCCATCCCAATTATGTGAAGCTGACACAGAC 545

RESULT 2 BD205037 8231 bp DNA linear PAT 17-JUL-2003
LOCUS Isolated SH3 gene relating to myeloproliferative disorders and
DEFINITION leukemia and utilization thereof.
ACCESSION BD205037

BD205037.1 GI:33014807
VERSION JP 2002511267-A/5.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Korenberg, J.R. and Chen, X.N.
JOURNAL 1 (bases 1 to 3231)
TITLE Isolated SH3 gene relating to myeloproliferative disorders and
leukemia and utilization thereof
REFERENCE Patent: JP 2002511267-A 5 16-APR-2002;
CEDARS SIVAI HEALTH SYSTEM ET AL
COMMENT OS Homo sapiens (human)
PN JP 2002511267-A/5
PD 16-APR-2002
PF 16-APR-1999 JP 2000543610
PR 16-APR-1998 US 60/082007
PI JULIE R KORENBERG, XIAO NING CHEN
PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,
PC C1201/68,
PC G01N33/68//A61K48/00,C12N15/00,C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and
leukemia
CC and utilization thereof.
FH Key Location/Qualifiers
FT source 1..3231
/organism="Homo sapiens (human)"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

Alignment Scores:
Pred. No.: 6.68e-32 Length: 3231
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x BD205037 (1-3231)

QY 1 AAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
Db 1730 GCACTGTGCCAGCGATTGGAGTGAAGACTACACCGCGCAAGATGACGATGAGCTGGCC 1789

QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTyrLysGly 40
Db 1790 TTCAACAAGGGCCGATCATCAACGCTCTCAACAAGGAGGACCTGACTGCTGAAAAGA 1849

QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59
Db 1850 GAAGTCATGACACAGTGGGCGCTTCCCATCCCAATTATGTGAAGCTGACACAGAC 1906

RESULT 3 HSU61166 3241 bp mRNA linear PRI 23-JUL-1996
LOCUS Human SH3 domain-containing protein SH3P17 mRNA, complete cds.
DEFINITION U61166
ACCESSION U61166.1 GI:1438932
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Sparks, A.B., Hoffman, N.G., McConnell, S.J., Fowlkes, D.M. and
KEY, B.K.
TITLE Cloning of ligand targets: systematic isolation of SH3
domain-containing proteins

JOURNAL	Nat. Biotechnol. 14 (6), 741-744 (1996)
MEDLINE	98294438
PUBMED	9630982
REFERENCE	2 (bases 1 to 3241)
AUTHORS	Pirozzi,G., McConnell,S.J., Uveges,A. and Fowlkes,D.M.
TITLE	Direct Substitution
JOURNAL	Submitted (18-JUN-1996) CYTOGEN Corp., 307 College Road East, Princeton, NJ 08540, USA
FEATURES	Location/Qualifiers
Source	1..3241
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/tissue_type="bone marrow"
	37..1599
CDS	/codon_start=1
	/product="SH3 domain-containing protein SH3P17"
	/protein_id="AAC0592.1"
	/db_xref="GI:1438933"
	/translation="MEAEFLKQKEQEKITIELEKQKEAQRPAERPKQMLHYEQED EHQKPRLHEEEKLREESVKKQDGEKRGQEADDKGR.LPHQDEPAKPAVQAPWST AEKPLRLISAQENVKVVYRALYPFESRSHDELITQGDIVNDSOTGPGWLGEL KKGKGFMPANYAEKIPENEEVAPVPAKADSTLSAPALRLRTPAPLAVTSBSTTP NNMADFSTWPTSTNEKEEDNMADPAQDSLTVPASGOLRORSAFPATATGSSEP VLQGGKEVEGIAQALYPMRAKDNHINPKNDVITVLBOODMMWPEVGOGKMPFK SYVLISGPIRKSTSMDSGSSSPASLSKRVASPAKAVVSSEELAQYATAYTAPPGQ LTTAPGQILIRKKNPGGMMWGEGLQAGKRGQIGMPFANYVKLSLPGTSKITTPERK STLAAAVCOVIGMYDYTAQNDDLELAFKNGQIINVINKEDPDWKKGVGVGLFPSNY VKLTITMDPSQ"
ORIGIN	
Alignment Scores:	
Pred. No.:	6.7e-32
Score:	323.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	9
Gaps:	0
US-09-720-934-2_COPY_1080_1138 (1-59) x HSU61166 (1-3241)	
QY	1 A1A1A1CysGlnValIleGlyMetTyrAspPyrThrAlaGlnAsnAspAspGluLeuAla 20
DB	1402 GCAGTGTCCACGAGTGATGGATGTGACACATCACCGCGAGATGACGATGAGCTGGCC 1461
QY	21 PheAsnTysGlyGlnIleIleAsnValLeuAsnTysGlnAspProAspTyrTyrGly 40
DB	1462 TTCAACAAGAGGCCACGATCATCACTCCTCAACAAGAGAGACCTGATGTGTGAAGA 1521
QY	41 GluValaGngIyGlnValGlyLeuPheProSerAsnTyrValIysLeuThrThrAsp 59
DB	1522 GAAGTCATGACAAAGTGGGGCTCTTCCATTCACATATATGTGAAGCTGACACAGAC 1578
RESULT 4	
LOCUS	HSN806384
DEFINITION	Homo sapiens mRNA, CDNA DKFZp686u17173 (from clone DKFZp686u17173).
ACCESSION	BX538175
VERSION	BX538175.1 GI:31874620
KEYWORDS	
SOURCE	
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 4321) Lauder,J.J., Bahr,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fodo,G., Han,M. and Wiemann,S.
TITLE	Direct Submision
JOURNAL	Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuberberg, GERMANY
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Sequenced by Olagen (Hilden/Germany) within the cDNA sequencing

consortium of the German Genome Project.
This clone (DKFZp666j1713) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cDNA/>.
Location/Qualifiers

1. .4321
/source="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="21q22.11"
/clone="DKFZp666j1713"
/tissue="human uterus endothel primary cell culture"
/clone_lib="686 (synonym: hlc63). Vector pSport1_Sfi; host
DH10B; sites Sfi1a + Sfi1b"
/dev_stage="adult"
1. .4321
/gene="DKFZp666j1713"
1. .1866
/gene="DKFZp666j1713"
/note="intersectin long isoform, N-terminus truncated,
differentially spliced"
/codon_start=3
/product="hypothetical protein"
/protein_id="CAD98050.1"
/db_xref="GI:31874621"
/translation="FFFFFLIVVHLLKAPVYOLPAPFAHPAATLPKSSSSRSBG
PGSOLNTRKQKQSPDVAVPVPAEMVNPSSRLKYLQLEFNSHDKTSGHLTGPQART
ILMOSLSPKOLASIVNLSIDIDQKLTAEFLIAMELILIVAMSGCLPVLPEEYIP
PSFPRVSGSGISVISTSDORPEEPVLEDEOOLEKLPVFEFKKRNFRGNIL
ELERRKALLFOCKREBERLAQLERAQOEKREPRQOEKROLEKLOKEBELRE
QREERKEIRERRAARELEKROQLDIERLTTGROEISTNKSELRIRAIETHLQOOLQSQQ
FEELANDKKHOLEKGLDIERLTTGROEISTNKSELRIRAIETHLQOOLQSQQ
LGRILPEKQILINDQLKQVQONSIRHDSIVLTKRALKEKELARQHLRDLQLEVEKETS
KLQSIDIPKQILNDKELRSTHNNKQOLKOKSMEAEILKQEOEKRIELEKQEZAKORRA
QERKQKLEHYVQOODEHORPRKLHEEKLRBESEVKKDKGEGKQKQEAODKLGLRPLHQ
HQBKRAVQAPMSTAEKGPLTISAENKAVYVYRALYPESRSHDETITGPQGIWVW
DES"
4267. .4272
/gene="DKFZp666j1713"
4289
/gene="DKFZp666j1713"

polyA_signal
4267. .4272
polyA_site
4289
/gene="DKFZp666j1713"

ORIGIN

Alignment Scores:

Pred. No.:	3.09e-32	Length:	4321
Score:	323.00	Matches:	59
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) x HSM806384 (1-4321)

QY 1 AAlaValCysGlnVal11leGlyMetCtyrAspTyrThrIaGlnAsnAspAspGluLeuAla 20

Db 2805 GCAGTGTGCCAGGCGATGGGATGTACACATCACCCGGCAGAAATGACGAGAGCTGGCC 2864

QY 21 PheaSnyVscGlyGlnIleIleAsnValLeuSnyGluuAspProAspTyrTrpIysGly 40

Db 2865 TTCAACAAGGGCCAGATATCAAGCTCTTCAACAGGAGGACCCCTGACTGGTGGAAAGA 2924

QY 41 GluValaengIyGlnValaIglyLeuPheProSerAsnTyrValIysLeuThrThrasp 59

Db 2925 GAAGTCATGACACAAGTGGGGCTCTTCCCATTCACATTATGTGCAAGCTACACACAGAC 2981

RESULT 5 BD205035 5195 bp DNA linear PAT 17-JUL-2003

LOCUS BD205035

DEFINITION Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.

ACCESSION BD205035

VERSION BD205035.1 GI:33014805
 KEYWORDS JP 2002511267-A/3.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 5195)
 KORENBERG, J.R. and Chen, X.N.
 Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof
 Parent: JP 2002511267-A 3 16-APR-2002;
 CEDARS SINAI HEALTH SYSTEM ET AL
 COMMENT OS Homo sapiens (human)
 PN JP 2002511267-A/3
 PD 16-APR-2002
 PF 16-APR-1999 JP 2000543610
 PR 16-APR-1998 US 60/082007
 PI JULIE R KORENBERG, XIAO NING CHEN
 PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10, C12Q1/68,
 PC G01N33/68//A61K48/00, C12N15/00, C12N5/00
 CC Isolated SH3 gene relating to myeloproliferative disorders and leukemia
 CC and utilization thereof.
 FH Key Location/Qualifiers
 FT source 1..5195
 FT location/Qualifiers
 1..5195
 /organism="Homo sapiens (human)"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1..1e-31 Length: 5195
 Score: 323.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_1080_1138 (1-59) x BD205035 (1-5195)
 QY 1 A1AvalCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
 Db 3689 GCAAGTGCACGAGGATGGAGTGTACGACTACACCGGCGAGAAATGACGATGAGCTGGCC 3748
 QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTrpLysGly 40
 Db 3749 TTCAACAAGAGGCGCATCATCATACCTCTCAACAAGAGGAGACCTGACTGTGGAAAGGA 3808
 QY 41 GluValAsnGlyGlnValIleGlyLeuPheProSerSerSeryValLysLeuThrThrasp 59
 Db 3809 GAAATCAATGACAAAGTGGGGCTCTCCATCCCAATTATGTGAAGCTGACCAAGAC 3865
 RESULT 6
 LOCUS BD205033 5199 bp DNA linear PAT 17-JUL-2003
 DEFINITION Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.
 ACCESSION BD205033
 VERSION BD205033.1 GI:33014803
 KEYWORDS JP 2002511267-A/1.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 5199)
 KORENBERG, J.R. and Chen, X.N.
 Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

JOURNAL Patent: JP 2002511267-A 1 16-APR-2002;
 CEDARS SINAI HEALTH SYSTEM ET AL
 COMMENT OS Homo sapiens (human)
 PN JP 2002511267-A/1
 PD 16-APR-2002
 PF 16-APR-1999 JP 2000543610
 PR 16-APR-1998 US 60/082007
 PI JULIE R KORENBERG, XIAO NING CHEN
 PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10, C12Q1/68,
 PC G01N33/68//A61K48/00, C12N15/00, C12N5/00
 CC Isolated SH3 gene relating to myeloproliferative disorders and leukemia
 CC and utilization thereof.
 FH Key Location/Qualifiers
 FT source 1..5199
 FT location/Qualifiers
 1..5199
 /organism="Homo sapiens (human)"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1..11e-31 Length: 5199
 Score: 323.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_1080_1138 (1-59) x BD205033 (1-5199)
 QY 1 A1AvalCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
 Db 3445 GCAAGTGCACGAGGATGGAGTGTACGACTACACCGGCGAGAAATGACGATGAGCTGGCC 3504
 QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTrpLysGly 40
 Db 3505 TTCAACAAGAGGCGCATCATCATACCTCTCAACAAGAGGAGACCTGACTGTGGAAAGGA 3564
 QY 41 GluValAsnGlyGlnValIleGlyLeuPheProSerSerSeryValLysLeuThrThrasp 59
 Db 3565 GAAATCAATGACAAAGTGGGGCTCTCCATCCCAATTATGTGAAGCTGACCAAGAC 3621
 RESULT 7
 LOCUS AF064243 5287 bp mRNA linear PRI 21-NOV-1998
 DEFINITION Homo sapiens interectin short form mRNA, complete cds.
 ACCESSION AF064243
 VERSION AF064243.1 GI:3859852
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 5287)
 GUIPPONI, M., Scott, H.S., Chen, H., Schebesta, A., Rossier, C. and Antonarakis, S.E.
 Two isoforms of a human interectin (ITSN) protein are produced by brain-specific alternative splicing in a stop codon
 JOURNAL GENOMICS 53 (3), 369-376 (1998)
 MEDLINE 99017974
 PUBMED 9799604
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Submitted (05-MAY-1998) Genetics and Microbiology, CMU, 1 rue Michel-Servet, Geneva 4 CH-1211, Switzerland
 Location/Qualifiers

```
source
1. 5287
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
/issue_type="brain"
/dev_stage="Fetus"
107.3769
/codon_start=1
/product="intersectin short form"
/protein_id="AAC78610.1"
/db_xref="GI:3659853"
/translation="MAQPTPPGGSLDIWAIVVERAKHDQGFHSIKPISGFTTGDOA
RNFPSGSLPQVLAQIWAIDANNNDGMDQVEFSIAMLIKIKLOGYQPSALPPLVM
KOQVATSSAPFPMGSIAMPELTVAVPWSIPIVGMSPITVSGVPLVAPPLAN
GAPVIOPLPAPFAPATLPKSSFSRSGPSQNLTKLOKQSFVAVPVAEMVP
OSSRLKYROLFNSHDKTMSGHLTGPOARTIMQSSLPQAKISITMWSLSDIDDGKLT
BEFLIAMHLIDVAMSGQPLPVLPPPIPPSFRVRSAGSISVISTSDQRLPEEPV
LDEQQQLERKLPVTFEDKKRENFERNLEKROALIEQKQERLAQRLAEQ
RKREERQERKQLELEKQLEKQREERKEIEREPAARELERQRLAE
RNRROELNQRNKEQEDIVLAKKKTLEFELEALNDKHOLEGKLDIRCHLTQRO
ETESTKSELELRIAEITHLQOOLQESQMLGRILPEKQILNDQKVOONSLSHDSLV
TLKRALAEKELARQHLRDLDEVEKETRSLQEDIDFNQIKELAREITHKQLOKXS
MEERLKQERKQLELEKQLEKQREERKEIEREPAARELERQRLAE
KRRESYKKGDEBKQERKQLEKQLEKQREERKEIEREPAARELERQRLAE
KVAVYRALYLPFSRSHDEITIOGDIVMKGWVDSQGEPMGLGELKGTGMEPA
NVAEKIPENVEVAPVAPVAPQSLTVPSAQLOKRSAPATATGSSPSVYLQGEVE
WPTSTNEKPEITNDMAAOPSLTVPSAQLOKRSAPATATGSSPSVYLQGEVE
GQAQALYPMRAKQNDHNLNKNVDITVLEQOMWFGVQGGKMPSPVYLISGP
IRKSTSDSGSSSPASLKRVAAPKPVSGEFLIANTYESSBEGDITFQGGVIL
VTKKSDMWTTGVGDAGVFPSPNYVLKDSBGSGLTAKTSLKKEIKQIVIASYAT
GEOQLTARQGLILIRKKNPGGWEGELQARGKQIGWFPANVYKLLNPGSKITPT
EPPKSTALAAVCQVIGMYDYTAQNDELAFNKGQILINVLNKEDPDMWKGEVNOQVGLF
PSNYVRLTTDMPESQ"
167.406
/misc_feature
/note="encodes EH domain"
767.1936
/misc_feature
/note="encodes EH domain"
2324.2524
/misc_feature
/note="encodes SH3 domain"
2843.3019
/misc_feature
/note="encodes SH3 domain"
3110.3286
/misc_feature
/note="encodes SH3 domain"
3326.3520
/misc_feature
/note="encodes SH3 domain"
3569.3748
/misc_feature
/note="encodes SH3 domain"
ORIGIN
Alignment Scores:
Pred. No.: 1,12e-31 Length: 5287
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-720-934-2_COPY_1080_1138 (1-59) x AF064243 (1-5287)
QY 1 AAlaValCyGlnValIleGlyMetYrTAspTyrThAlaGlnAsnAspAspGluLeuAla 20
Db 3572 GCAgTGTGCCAGGTGATGGATGTACGACTACACCGCGAATGACGATGAGCTGGCC 3631
QY 21 PheAsnlysglyGlnIleIleAsnValIleuAsnlysglyuAspProAspTyrTPlYsgly 40
Db 3632 TTtAAcAAAGGCGCAGATCATCAACGTCCTCAACAAGAGAGACCCTGATCGTGGAAAGCA 3691
QY 41 GluValAsnlyGlnValGlyLeuPheProSerAsnTyrVallyLeuThrThrAsp 59
Db 3692 GAAGTCAATGAGCAAGTGGGGCTCTTCCCATCAATTATGTGAAGCTGACCAAGAC 3748
```

```
RESULT 8
AF114488
LOCUS AF114488 5381 bp mRNA linear PRI 16-JUL-2002
DEFINITION Homo sapiens intersectin short isoform (ITSN) mRNA, complete cds.
ACCESSION AF114488
VERSION AF114488.1 GI:4808824
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 5381)
Pucharcos,C., Fuentes,J.J., Pritchard,M. and Estivill,X.
Alu-splice cloning of human intersectin (ITSN), a putative
multivalent binding protein expressed in proliferating and
differentiating neurons and overexpressed in Down syndrome
Eur. J. Hum. Genet. 7 (6), 704-712 (1999)
JOURNAL
MEDLINE 99415290
PUBMED 10482960
REFERENCE
2 (bases 1 to 5381)
Pucharcos,C., Fuentes,J.J., Pritchard,M. and Estivill,X.
Direct Submition
Submitted (16-DEC-1998) Medical and Molecular Genetics Center,
Cancer Research Institute, L'Hospitalet de llo., Avia.
Castelldefels Km. 2,7, Barcelona 08907, Spain
Location/Qualifiers
1. 5381
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
1.5381
/gene="ITSN"
269.3931
/gene="ITSN"
/codon_start=1
/product="Intersectin short isoform"
/protein_id="AAd29953.1"
/db_xref="GI:4808825"
/translation="MAQPTPPGGSLDIWAIVVERAKHDQGFHSIKPISGFTTGDOA
RNFPSGSLPQVLAQIWAIDANNNDGMDQVEFSIAMLIKIKLOGYQPSALPPLVM
KOQVATSSAPFPMGSIAMPELTVAVPWSIPIVGMSPITVSGVPLVAPPLAN
GAPVIOPLPAPFAPATLPKSSFSRSGPSQNLTKLOKQSFVAVPVAEMVP
OSSRLKYROLFNSHDKTMSGHLTGPOARTIMQSSLPQAKISITMWSLSDIDDGKLT
BEFLIAMHLIDVAMSGQPLPVLPPPIPPSFRVRSAGSISVISTSDQRLPEEPV
LDEQQQLERKLPVTFEDKKRENFERNLEKROALIEQKQERLAQRLAEQ
RKREERQERKQLELEKQLEKQREERKEIEREPAARELERQRLAE
KRRESYKKGDEBKQERKQLEKQLEKQREERKEIEREPAARELERQRLAE
KVAVYRALYLPFSRSHDEITIOGDIVMKGWVDSQGEPMGLGELKGTGMEPA
NVAEKIPENVEVAPVAPVAPQSLTVPSAQLOKRSAPATATGSSPSVYLQGEVE
WPTSTNEKPEITNDMAAOPSLTVPSAQLOKRSAPATATGSSPSVYLQGEVE
GQAQALYPMRAKQNDHNLNKNVDITVLEQOMWFGVQGGKMPSPVYLISGP
IRKSTSDSGSSSPASLKRVAAPKPVSGEFLIANTYESSBEGDITFQGGVIL
VTKKSDMWTTGVGDAGVFPSPNYVLKDSBGSGLTAKTSLKKEIKQIVIASYAT
GEOQLTARQGLILIRKKNPGGWEGELQARGKQIGWFPANVYKLLNPGSKITPT
EPPKSTALAAVCQVIGMYDYTAQNDELAFNKGQILINVLNKEDPDMWKGEVNOQVGLF
PSNYVRLTTDMPESQ"
ORIGIN
Alignment Scores:
Pred. No.: 1,15e-31 Length: 5381
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
```

US-09-720-934-2_COPY_1080_1138 (1-59) x AF114448 (1-5381)	
QY 1 A1aValCySGInVal111eGlyMeTyrAspTyrThrAlaGlnAsnAspAspGlnLeuA1a 20	
Db 3734 GCAGGTGCCAGGATGGATGGATGTACGACTACACCGCGCAGAAATGACGATGAGCTGGCC 3793	
QY 21 PheAsnYsG1yGln11e1leAsnValLeuAsnYsG1uAspProAspTyrTrpYsG1y 40	
Db 3794 TTCAACAAGGGCCAGATATCATCAACGTCCTCTCAACAAGAGAGACCCCTGACTGGTGGAAAGGA 3853	
QY 41 GlnValaAsnG1yGlnValaG1yLeuPheProSerAsnTyrValYsLeuThrThrAsp 59	
Db 3854 GAAGTCATGAGCAAGGTGGGGGTCTTCCCATTCACATTTATGTGAAGCTACACCAACAC 3910	
RESULT 9	
LOCUS BD205034	5458 bp DNA linear PAT 17-JUL-2003
DEFINITION Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof.	
ACCESSION BD205034	
VERSION BD205034.1 GI:33014804	
KEYWORDS JP 2002511267-A/2.	
SOURCE Homo sapiens (human)	
ORGANISM Homo sapiens	
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS 1 (bases 1 to 5458)	
TITLE Korenberg,J.R. and Chen,X.N.	
JOURNAL Isolated SH3 gene relating to myeloproliferative disorders and leukemia and utilization thereof	
COMMENT Patent: JP 2002511267-A 2 16-APR-2002;	
CEPDA SIVAI HEALTH SYSTEM ET AL	
OS Homo sapiens (human)	
PN JP 2002511267-A/2	
PD 16-APR-2002	
PF 16-APR-1999 JP 2000B543610	
PR 16-APR-1998 US 60/082007	
PI JULIE R. KORENBERG, XIAO NING CHEN	
PC C12N15/09,A01K67/027,C07K14/47,C07K16/18,C07K19/00,C12N5/10,	
PC C12Q1/68,	
PC G01N33/68//A61K48/00,C12N15/00,C12N5/00	
CC Isolated SH3 gene relating to myeloproliferative disorders and	
CC CC and utilization thereof.	
CC FH Key Location/Qualifiers	
FT source 1..5458	
FT Location/Qualifiers	
source 1..5458 /organism='Homo sapiens (human)'. /location='chromosome 17p11.2'. /db_xref='taxon:9606'	
ORIGIN	
Alignment Scores:	
Pred. No.:	1.16e-31
Score:	323.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	6
US-09-720-934-2_COPY_1080_1138 (1-59) x BD205034 (1-5458)	
QY 1 A1aValCySGInVal111eGlyMeTyrAspTyrThrAlaGlnAsnAspAspGlnLeuA1a 20	
Db 3732 GCAGGTGCCAGGATGGATGGATGTACGACTACACCGCGCAGAAATGACGATGAGCTGGCC 3791	
QY 21 PheAsnYsG1yGln11e1leAsnValLeuAsnYsG1uAspProAspTyrTrpYsG1y 40	
Db 3792 TTCAACAAGGGCCAGATATCATCAACGTCCTCTCAACAAGAGAGACCCCTGACTGGTGGAAAGGA 3851	

[illegible]

ORIGIN

CRSHGKSNPYCEVTWMSQCHITKTIDTILNPKMNSCOFFIRDLBOELCTIVFERDQ
FSPDDELGRTEIRIVADIKKDQSGKGVTKCLLHHEVPTGSEIVRDLQLFDEB"

Alignment Scores:

Pred. No.:	1.39e-31	Length:	6439
Score:	323.00	Matches:	59
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) x AF114487 (1-6439)

QY 1 AlAlaValCysGlnValIleGlyMetTyTAspTyThrAlaGlnAsnAspAspGluLeuAla 20
Db 3734 GCAGTGTGCCAGGTGATGGATGTACGACTACACCCGCGAGATGACGTAGCTGGGCC 3793
QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTPTPTLysGly 40
Db 3794 TTCAACAAGGGCCAGATCATCAACGTCCTCAACAAGAGGAGCCCTGACCTGTGAAAGGA 3853
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyTValLysLeuThrThrAsp 59
Db 3854 GAAGTCAATGACCAAGTGGGGCTCTTCCCATCCATTAATGTAAGCTGACCAAGAC 3910

RESULT 11
AF064244 7247 bp mRNA linear PRI 21-NOV-1998

LOCUS Homo sapiens intersecin long form mRNA, complete cds.
DEFINITION AF064244
ACCESSION AF064244
VERSION AF064244.1 GI:3859854

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia: Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 7247)
Guidpomi, M., Scott, H. S., Chen, H., Schebesta, A., Rossier, C. and Antonaakis, S. B.

TITLE
JOURNAL 2 (bases 1 to 7247)
MEDLINE 99017974
PUBMED 9799604
REFERENCE 2 (bases 1 to 7247)
AUTHORS Guidpomi, M., Scott, H. S., Chen, H., Schebesta, A., Rossier, C. and Antonaakis, S. B.

TITLE
JOURNAL Submitted (05-MAY-1998) Genetics and Microbiology, CMU, 1 rue Michel-Servet, Geneva 4 CH-1211, Switzerland
FEATURES
source Location/Qualifiers
1. 7247
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
/tissue_type="brain"
/dev_stage="fetus"
107..5272
/codon_start=1
/product="intersecin long form"
/protein_id="AAC78611.1"
/db_xref="GI:3859854"
/translation="MAQPTPPFGSLDIWATVEERAKHDQPHSKPIISGFIITGQAR
RNPFOSGLPOPLAIOIWLADNNDNDMOVCFSTAMKLIKLOGLPSALPEVM
KOOPVAISAPRPGMGIAAMPBILTAAPVPMGSIIVMGSPITVSSVPTAAAPLAP
GAPVIOPLPAPAPPAATLPIKSSFSRSGSLNKLQKAGSFVDSAPVAPKAVP
QSERLKRQLPNSHDKTMSHLTGPQKRTILMSSLPQALASIMLSDIDQKULTA
BEFTLAMHLIDVAMSGQPLPVPILPEYIPSPFSFRVSSGSIIVSISVSDQRLPEEP
LEDEQOOLEKRLPTEFEDKKRENERGNTLEKRRQALLQQKREORLAQLRAECS
RKRREROEOERKKOLELEKQLELREQRREBERREKREIRREARRELEKQQLME

CDS

US-09-720-934-2_COPY_1080_1138 (1-59) x AF064244 (1-7247)

ORIGIN

Alignment Scores:
Pred. No.: 1.57e-31
Score: 323.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 9

QY 1 AlAlaValCysGlnValIleGlyMetTyTAspTyThrAlaGlnAsnAspAspGluLeuAla 20
Db 3572 GCAGTGTGCCAGGTGATGGATGTACGACTACACCCGCGAGATGACGTAGCTGGGCC 3631
QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTPTPTLysGly 40
Db 3632 TTCAACAAGGGCCAGATCATCAACGTCCTCAACAAGAGGAGCCCTGACCTGTGAAAGGA 3691
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyTValLysLeuThrThrAsp 59
Db 3692 GAAGTCAATGACCAAGTGGGGCTCTTCCCATCCATTAATGTAAGCTGACCAAGAC 3748

RESULT 12
AP000312 12015 bp DNA linear PRI 24-MAY-2003
LOCUS Homo sapiens genomic DNA, chromosome 21 clone:RP1-201f12, complete
DEFINITION sequence.
ACCESSION AP000312
VERSION AP000312.2 GI:31071646

KEYWORDS	HTG.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Tokoki,Y., Watanabe,H. and Sakaki,Y.
TITLE	2 Published Only in Database (1999)
JOURNAL	3 (bases 1 to 12015)
REFERENCE	4 Homo sapiens genomic DNA
AUTHORS	5 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Tokoki,Y., Watanabe,H. and Sakaki,Y.
TITLE	6 Direct Submission
JOURNAL	7 Submitted (13-MAY-1999) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT	8 On May 23, 2003 this sequence version replaced gi:4835681.
FEATURES	9 Location/Qualifiers
SOURCE	10 1..12015
ORIGIN	11 /organism="Homo sapiens"
	12 /mol_type="genomic DNA"
	13 /db_xref="taxon:9606"
	14 /chromosome="21"
	15 /map="21q"
	16 /clone="RP1-201F12"
Alignment Scores:	17
Prd. No.:	18 2.68e-31
Score:	19 323.00
Percent Similarity:	20 100.00%
Best local Similarity:	21 100.00%
Query Match:	22 100.00%
DB:	23 9
	24 Gaps: 0
US-09-720-934-2_COPY_1080_1138 (1-59) x AP000312 (1-12015)	
OY	25 1 AAlaATCGsGlnValIleGlyMeCTyrAspTyrThrIaGlnAsnAspAspGluLeuA 20
DB	26 2652 GCAATGTCGCCAGCGATGGATGGATGACACATCACCGCGCAATGACGATGAGCTGCC 27111
OY	27 21 PheAnlnysGlyGlnIleIleAsnValleuAnlnysGluAspProAspTyrTyrIly 40
DB	28 2712 TTCAACAAGAGCGCCAGATCATCAACGCTCCTCAACAAGAGGAGCACCCTGACTGGAAAGGA 27711
OY	29 41 GlnValasnGlyGlnValGlyIleuPheProSerAsnTyrValIysIleuThrThrasp 59
DB	30 2772 GAAGTCATGACGACAAAGTGCGGCTCTCCCATCCCAATTATGTGAAGCTGACCAAGAC 2828
RESULT 13	
AP000193	100000 bp DNA linear PRI 08-JAN-2000
LOCUS	Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region, clone Q78C10-F32B9, segment 20/21, complete sequence.
ACCESSION	AP000193
VERSION	AP000193.1 GI:4826584
KEYWORDS	HTG.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Tokoki,Y., Watanabe,H. and Sakaki,Y.
TITLE	2 Published Only in Database (1999)
JOURNAL	3 (bases 1 to 100000)
REFERENCE	4 Homo sapiens genomic DNA of 21q22.1 (REGION: D21S226-AML CLONE RANGE: Q78C10-F32B9)
AUTHORS	5 Published Only in Database (1999)
TITLE	6 (bases 1 to 100000)
JOURNAL	7 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Tokoki,Y. and Sakaki,Y.
REFERENCE	8 Homo sapiens genomic DNA of 21q22.1 (REGION: D21S226-AML CLONE RANGE: Q78C10-F32B9)
AUTHORS	9 Published Only in Database (1999)
TITLE	10 (bases 1 to 100000)
JOURNAL	11 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Tokoki,Y. and Sakaki,Y.

COMMENT	FEATURES	ORIGIN	ALIGNMENT SCORES	US-09-720-934-2_COPY_1080_1138 (1-59) x AP000193 (1-100000)
JOURNAL	Submitted (10-MAY-1999) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan (E-mail:hattori@gsc.riken.go.jp), URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923, Fax:81-42-778-9924)			
COMMENT	E. coli transposon insertion: The present data does not contain E. coli transposon sequences which integrated in the original/previous sequences. We determined the boundary between the insertion and genomic sequences experimentally, removed the insertion sequences, reconstituted the present data. The sequencing project is supported by Japan Science Technology Corporation (JST) and The Institute of Physical and Chemical Research (RIKEN).			
SOURCE	Location/Qualifiers 1. 100000 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="21" /map="21q22.1"			
ALIGNMENT SCORES				
Pred. No.:	2,52e-30	Length:	100000	
Score:	323.00	Matches:	59	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	100.00%	Indels:	0	
DB:	9	Gaps:	0	
QY	1	AlaValysGlnValIleGlyMetTyrAspTyrThrIaGlnAsnAspGluLeuAla	20	
DB	64215	GCACTGTGCCACAGTGGATTGGAGTGTACGACTACACCGCGAGATGACGTGGCTGGCC	64227	
QY	21	PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTyrGlyGly	40	
DB	64275	TTCAACACAGGCGCCGACATCAACGTCCTCCACACAGAGAGACCTGACTGGTGGAAAGGA	64333	
QY	41	GluValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp	59	
DB	64335	GAAATCAATGACACAGTGGGCGCTCTTCCATCCAAATTATGTGAAGCTGACACACAGAC	64391	
RESULT 14				
AP000050	114929 bp	DNA	linear	PRI 20-NOV-1999
LOCUS	Homo sapiens genomic DNA, chromosome 21q22.1, segment 21/28,			
DEFINITION	complete sequence.			
ACCESSION	AP000050			
VERSION	AP000050.1			
KEYWORDS	HTG.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
AUTHORS	1 (bases 1 to 114929)			
TITLE	Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.			
JOURNAL	Homo sapiens genomic DNA, chromosome 21q			
REFERENCE	Published Only in Database (1998)			
AUTHORS	2 (bases 1 to 114929)			
TITLE	Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.			
JOURNAL	Direct Submission			
COMMENT	Submitted (11-MAY-1998) Masahira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan (E-mail:hattori@gsc.ims.u-tokyo.ac.jp), Tel:0427-78-9733, Fax:0427-78-9561)			
	This sequence is conducted by Kitasato University JST sequencing Laboratory as a JST sequencing team.			
	Principal Investigator: Yoshiyuki Sakaki Ph.D.			
	Phone:+81-3-5449-5622, Fax : +81-3-5449-5445,			
	sakaki@hgc.ims.u-tokyo.ac.jp			

Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 9

Mismatches: 0
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x AP000117 (1-151516)

QY	1	AlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla	20
DB	64215	GCAGTGTCCCGAGTGTATGGGATGTACACTACACCGCGCAGAAATGACGATGAGCTGSCC	64274
QY	21	PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGlyAspProAspTyrTrpLysGly	40
DB	64275	TTCAACAAGGCGCCAGATCATCAAGCTCTCAACAAGAGAGACCCCTGACTGGTGGAAAGGA	64334
QY	41	GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrasp	59
DB	64335	GAAGTCAATGAGACAGTGGGCTCTTCCCATCAATTATGTGAAGTGAACCAACAGAC	64391

Search completed: July 1, 2004, 16:16:09
Job time : 1391.81 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:03:00 ; Search time 41.318 Seconds

(without alignments)
792.522 Million cell updates/sec

Title: US-09-720-934-2_COPY_1080_1138
Sequence: 323
1 AVOQVIGMYDYAQNDELIA.....GEVNGGVLPSPNYVLTLD 59

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_p2n.model -DBV=xlp
-Q=/cgn2_1/USFTO_spool_p/US09720934/rnuc 30062004_064540_13455/app_query.fasta_1.1386
-DB=Issued Patents NA -QFMT=fastap -SUPFIX=rml -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=humana40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=0.5
-MODE=LOCAL -OUTFMT=pico -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09720934@cgn1_1.284@runcat 30062004_064540_13455 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOUTER -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEROUT=120 -WARN_TIMEROUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6C.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/6D.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	266	82.4	747	4	US-08-630-915A-39
2	266	82.4	2873	4	US-08-630-915A-193
3	160	49.5	1392	1	US-08-475-894-3
4	160	49.5	1392	1	US-08-484-710-3
5	160	49.5	1392	2	US-08-484-710-3
6	160	49.5	1392	2	US-08-484-710-3
7	160	49.5	1392	2	US-08-484-710-3
8	160	49.5	1392	2	US-08-484-710-3
9	160	49.5	1392	2	US-08-484-710-3
10	160	49.5	1392	2	US-08-484-710-3
11	160	49.5	1392	2	US-08-484-710-3
12	156	48.3	1045	1	US-08-475-894-5

13	156	48.3	1045	1	US-08-484-710-5	Sequence 5, Appl
14	156	48.3	1045	2	US-08-484-709-5	Sequence 5, Appl
15	156	48.3	1045	3	US-08-474-697-5	Sequence 5, Appl
16	154	47.7	1430	2	US-08-549-004A-15	Sequence 15, Appl
17	154	47.7	1430	3	US-09-051-982A-15	Sequence 15, Appl
18	152	47.1	813	4	US-09-023-655-1238	Sequence 138, Ap
19	152	47.1	933	1	US-09-641-640-1	Sequence 1, Appl
20	152	47.1	933	2	US-08-612-857-1	Sequence 1, Appl
21	152	47.1	984	4	US-09-765-298A-5	Sequence 5, Appl
22	152	47.1	1072	1	US-07-906-349A-2	Sequence 2, Appl
23	152	47.1	1072	1	US-08-167-035-5	Sequence 5, Appl
24	152	47.1	1072	1	US-08-167-035-49	Sequence 49, Appl
25	152	47.1	1072	1	US-08-208-887A-5	Sequence 5, Appl
26	152	47.1	1072	2	US-08-539-005-49	Sequence 5, Appl
27	152	47.1	1072	2	US-08-539-005-49	Sequence 49, Appl
28	152	47.1	1072	2	US-08-539-005-49	Sequence 2, Appl
29	152	47.1	1109	3	US-08-664-962B-5	Sequence 2, Appl
30	152	47.1	1109	3	US-09-311-743-5	Sequence 5, Appl
31	150	46.4	1661	2	US-08-815-176-2	Sequence 2, Appl
32	150	46.4	1661	2	US-08-815-176-2	Sequence 2, Appl
33	148	45.8	1710	4	US-08-630-915A-189	Sequence 189, App
34	144	44.6	1640	4	US-09-023-655-650	Sequence 650, App
35	143.5	44.4	1347	4	US-09-833-381-1937	Sequence 931, App
36	142	44.0	874	4	US-08-942-423-68	Sequence 68, Appl
37	142	44.0	1458	2	US-08-942-423-68	Sequence 25, Appl
38	142	44.0	2003	4	US-08-942-423-1	Sequence 32, Appl
39	142	44.0	2003	4	US-08-942-423-1	Sequence 32, Appl
40	141	43.7	949	1	US-08-167-035-32	Sequence 32, Appl
41	141	43.7	949	1	US-08-167-035-50	Sequence 50, Appl
42	141	43.7	949	1	US-08-208-887A-32	Sequence 32, Appl
43	141	43.7	949	2	US-08-539-005-32	Sequence 50, Appl
44	141	43.7	949	2	US-08-539-005-50	Sequence 34, Appl
45	141	43.7	949	4	US-09-280-598-34	

ALIGNMENTS

RESULT 1
US-08-630-915A-39
; Sequence 39, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOMLICKS, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Misticok, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 747 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-630-915A-39

Alignment Scores:
Pred. No.: 5.57e-30
Score: 266.00
Percent Similarity: 91.38%
Best Local Similarity: 75.86%
Query Match: 82.35%
DB: 4

Length: 747
Matches: 44
Conservative: 9
Mismatch: 5
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-630-915A-39 (1-747)

QY 2 ValCysGlnValIleGlyMetTyrAspTyrThrIaGlnAsnAspAspGlnLeuAlaPhe 21
Db 553 GTATGTCAGGTGATGCTATGATGACTATGACCAATAATATGAAGATGAGCTCAGTTTC 612

QY 22 AsnlysglyGlnIlelleAsnValLeuAsnlysglnAspProAspTyrTrpTyrlyGlyGlu 41
Db 613 TCCAAGGACACATTAATGATTATGAACAAAGATGCTGATGTTGGTGCAGAGAG 672

QY 42 ValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrVallyLeuThrThrAsp 59
Db 673 ATCAACGGGGGAGCTGCTCTTCTTCCTTCAACACTACGTTAAGATGACGACAGAC 726

RESULT 2
US-08-630-915A-193
Sequence 193, Application US/08630915A
Patent No. 6309820
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLES, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
NUMBER OF INVENTION: USING SAME
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mastrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 193:
SEQUENCE CHARACTERISTICS:

LENGTH: 2873 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-08-630-915A-193

Alignment Scores:
Pred. No.: 3.08e-29
Score: 266.00
Percent Similarity: 91.38%
Best Local Similarity: 75.86%
Query Match: 82.35%
DB: 4

Length: 2873
Matches: 44
Conservative: 9
Mismatch: 5
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-630-915A-193 (1-2873)

QY 2 ValCysGlnValIleGlyMetTyrAspTyrThrIaGlnAsnAspAspGlnLeuAlaPhe 21
Db 1338 GTATGTCAGGTGATGCTATGATGACTATGACCAATAATATGAAGATGAGCTCAGTTTC 1397

QY 22 AsnlysglyGlnIlelleAsnValLeuAsnlysglnAspProAspTyrTrpTyrlyGlyGlu 41
Db 1398 TCCAAGGACACATTAATGATTATGAACAAAGATGCTGATGTTGGTGCAGAGAG 1457

QY 42 ValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrVallyLeuThrThrAsp 59
Db 1458 ATCAACGGGGGAGCTGCTCTTCTTCCTTCAACACTACGTTAAGATGACGACAGAC 1511

RESULT 3
US-08-475-894-3
Sequence 3, Application US/08475894
Patent No. 5641748
GENERAL INFORMATION:
APPLICANT: Yen-Ming Hsu
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,894
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-191
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1392 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-475-894-3

Alignment Scores:
Pred. No.: 4.16e-14
Score: 160.00
Percent Similarity: 73.21%
Length: 1392
Matches: 29
Conservative: 12

Page 3

Db 81.7 GAGGAGATATAGTCACCTTCATCATTAAGACTGCATCGACTGCTGGGAAGA 876

[illegible]

RESULT 5
US-08-484-709-3

Patent No. 5837844
GENERAL INFORMATION:
APPLICANT: Yen-Ming Hsu

```

; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
;

```

CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

```
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,709
 FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965

```

; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: (617) 227-7400
;
; TELEFAX: (617) 227-5941
;
; TRANSMISSION FOR: 227-5941

```

```

;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 1392 base pairs
; TYPE: nucleic acid
;
;

```

```

;          TOPOLOGY: linear
;          MOLECULE TYPE: CDNA
US-08-484-709-3

```

Alignment Scores:	
Pred. No.:	4.16e-14
Score:	Length: 1392
	Matches: 29
Percent similarity:	2
Difference:	10

Best Local Similarity:	51.79%	Mismatches:	11
Query Match:	49.54%	Indels:	4
DB:	2	Gaps:	2

QY 3 CysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuA
US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-484-709-3 (1-1392)

Db 763 TGC A A G T A T A ----- T T C C A T T G A G G C A C A G A T G A T G A T T G A
Qy 23 L y S G l y G i n I l e I e a s n V a l l e u s n l y S G i u ----- A s p P r o a s p T r p T

Db 817 GAAGGAGATATAGTCACTCTCATCATAAGGACTGCATCGACGTAGCGTGGT

DB 8/7 GAGCTGAACGGCAGACGAGCGCTGTTCCCGATAACTTCGTGAAGTTA 924

; sequence 3, Application US/084/4697
; Patent No. 6171800

```

GENERAL INFORMATION:
APPLICANT: Yen-Ming Hsu
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,697
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-193
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1392 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-474-697-3

Alignment Scores:
Pred. No.: 4.15e-14      Length: 1392
Score: 160.00           Matches: 29
Percent Similarity: 73.21% Conservative: 12
Best Local Similarity: 51.79% Mismatches: 11
Query Match: 49.54%      Indels: 4
Gaps: 2

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-474-697-3 (1-1392)
QY 3 CysGlnVal11IeglyMetYrAspTyThrAlaGlnAsnAspAspGluLeuAlaPheAsn 22
Db 763 TGCAGAGTAATA-----TTTCCATATGAGGACACAGATGATGATGATGACATCAAA 816
QY 23 LysGlyGln11IleAsnValLeuAsnLysGlu-----AappProAspTyrTrpLysGly 40
Db 817 GAAGGAGATATAGTACTCTCATCAATAGAGACTGCATCGACGTRAGGCTGGTGGAGAGA 876
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyValLysLeu 56
Db 877 GAGCTGAACGCGACAGCGAGCGGTGTTCCCGATTAACCTCGGAAGTTA 924

RESULT 7
US-08-475-894-1
; Sequence 1, Application US/08475894
; Patent No. 5641748
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,894
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-191
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1659 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-475-894-1

Alignment Scores:
Pred. No.: 5.19e-14      Length: 1659
Score: 160.00           Matches: 29
Percent Similarity: 73.21% Conservative: 12
Best Local Similarity: 51.79% Mismatches: 11
Query Match: 49.54%      Indels: 4
Gaps: 2

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-475-894-1 (1-1659)
QY 3 CysGlnVal11IeglyMetYrAspTyThrAlaGlnAsnAspAspGluLeuAlaPheAsn 22
Db 1030 TGCAGAGTAATA-----TTTCCATATGAGGACACAGATGATGATGATGACATCAAA 1083
QY 23 LysGlyGln11IleAsnValLeuAsnLysGlu-----AappProAspTyrTrpLysGly 40
Db 1084 GAAGGAGATATAGTACTCTCATCAATAGAGACTGCATCGACGTRAGGCTGGTGGAGAGA 1143
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyValLysLeu 56
Db 1144 GAGCTGAACGCGACAGCGAGCGGTGTTCCCGATTAACCTCGGAAGTTA 1191

RESULT 8
US-08-484-710-1
; Sequence 1, Application US/08484710
; Patent No. 5656438
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,710
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
```

```
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1659 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-484-710-1

Alignment Scores:
Pred. No.: 5.19e-14 Length: 1659
Score: 160.00 Matches: 29
Percent Similarity: 73.21% Conservative: 12
Best Local Similarity: 51.79% Mismatches: 11
Query Match: 49.54% Indels: 4
DB: 1 Gaps: 2

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-484-710-1 (1-1659)

QY 3 CysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGlnLeuAlaPheAsn 22
Db 1030 TCGAAGTATA-----TTCCATATGAGGCACAGATGATGATGATGACATCAAA 1083

QY 23 LysGlyGlnIleLeuAsnValLeuAsnLysGlu-----AspProAspTyrPrlPylsGly 40
Db 1084 GAAGGAGATATAGTACTCATCATCAATAGAGACTGCATCGAGCTGAGCTGGGAAGA 1143

QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeu 56
Db 1144 GAGCTGAACGACAGACAGCGCGTTCCTCCGATTAACCTTCGTGAAGTTA 1191

RESULT 9
US-08-484-709-1
Sequence 1, Application US/08484709
Patent No. 5837844
GENERAL INFORMATION:
APPLICANT: Yen-Ming Hsu
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,709
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-192
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1659 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
```

```
MOLECULE TYPE: cDNA
US-08-484-709-1

Alignment Scores:
Pred. No.: 5.19e-14 Length: 1659
Score: 160.00 Matches: 29
Percent Similarity: 73.21% Conservative: 12
Best Local Similarity: 51.79% Mismatches: 11
Query Match: 49.54% Indels: 4
DB: 2 Gaps: 2

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-484-709-1 (1-1659)

QY 3 CysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGlnLeuAlaPheAsn 22
Db 1030 TCGAAGTATA-----TTCCATATGAGGCACAGATGATGATGATGACATCAAA 1083

QY 23 LysGlyGlnIleLeuAsnValLeuAsnLysGlu-----AspProAspTyrPrlPylsGly 40
Db 1084 GAAGGAGATATAGTACTCATCATCAATAGAGACTGCATCGAGCTGAGCTGGGAAGA 1143

QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeu 56
Db 1144 GAGCTGAACGACAGACAGCGCGTTCCTCCGATTAACCTTCGTGAAGTTA 1191

RESULT 10
US-08-474-697-1
Sequence 1, Application US/08474697
Patent No. 6171800
GENERAL INFORMATION:
APPLICANT: Yen-Ming Hsu
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,697
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-193
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1659 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-474-697-1

Alignment Scores:
Pred. No.: 5.19e-14 Length: 1659
Score: 160.00 Matches: 29
Percent Similarity: 73.21% Conservative: 12
Best Local Similarity: 51.79% Mismatches: 11
Query Match: 49.54% Indels: 4
DB: 3 Gaps: 2
```

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-474-697-1 (1-1659)

QY 3 CysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspArgGluLeuAlaPheAsn 22
DB 1030 TGCAGAGTAATA-----TTTCCATATAGAGCAGAGATGATGATGATTGACATCAAA 1083
QY 23 LysGlyGlnIleIleAsnValIleuAsnLysGlu-----AspProAspTyrTrpLysGly 40
DB 1084 GAGGAGATATAGCACTCTCATCAATAGAGACTGCATCGACGTAAGCTGGTGGGAAGA 1143
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeu 56
DB 1144 GAGCTGAACGCGACAGACGAGCGTGTCCCGATTAATTCTGGAAGTTA 1191

RESULT 11
US-08-671-354-1
; Sequence 1, Application US/08671354
; Patent No. 6423824
; GENERAL INFORMATION:
; APPLICANT: Hsu, Yen-Ming
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/671.354
; FILING DATE: 27-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,344
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/484,709
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/475,894
; FILING DATE: 07-JAN-1995
; APPLICATION NUMBER: US 08/475,710
; FILING DATE: 07-JAN-1995
; APPLICATION NUMBER: US 08/474,697
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10274/009005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3143 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 436...2412
; US-08-671-354-1

Alignment Scores:
Pred. No.: 1.17e-13
Score: 160.00
Percent Similarity: 73.21%
Best Local Similarity: 51.79%

Length: 3143
Matches: 29
Conservative: 12
Mismatch: 11

Query Match: 49.54% Indels: 4
DB: 4 Gaps: 2

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-671-354-1 (1-3143)

QY 3 CysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspArgGluLeuAlaPheAsn 22
DB 1231 TGCAGAGTAATA-----TTTCCATATAGAGCAGAGATGATGATGATTGACATCAAA 1284
QY 23 LysGlyGlnIleIleAsnValIleuAsnLysGlu-----AspProAspTyrTrpLysGly 40
DB 1285 GAGGAGATATAGCACTCTCATCAATAGAGACTGCATCGACGTAAGCTGGTGGGAAGA 1344
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeu 56
DB 1345 GAGCTGAACGCGACAGACGAGCGTGTCCCGATTAATTCTGGAAGTTA 1392

RESULT 12
US-08-475-894-5
; Sequence 5, Application US/08475894
; Patent No. 5641748
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,894
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-191
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1045 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-475-894-5

Alignment Scores:
Pred. No.: 1.11e-13 Length: 1045
Score: 156.00 Matches: 26
Percent Similarity: 73.08% Conservative: 12
Best Local Similarity: 50.00% Mismatches: 14
Query Match: 48.30% Indels: 0
DB: 1 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-475-894-5 (1-1045)

QY 4 GluValIleGlyMetTyrAspTyrThrAlaGlnAsnAspArgGluLeuAlaPheAsnLys 23
DB 118 GAGGCAATAGTGAGATTGACTACAGGCCAGACGATGATGATGACATCGCGTG 177
QY 24 GlyGlnIleIleAsnValIleuAsnLysGluAspProAspTyrTrpLysGlyGluValAsn 43

Db 178 GGTGAATCATCAACCAATCAGGAAGAGATGAGCTGTGGAGGACAGATCAAC 237
QY 44 G|yG|nValG|yLeuPheProSeRAsnTyValLys 55
Db 238 GGCAGAGAGGTTTGTTCCTGACCAACTTGTAAAG 273

RESULT 13

US-08-484-710-5
; Sequence 5, Application US/08484710
; Patent No. 5636438
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,710
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1045 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-484-710-5

Alignment Scores:
Pred. No.: 1,11e-13 Length: 1045
Score: 156.00 Matches: 26
Percent Similarity: 73.08% Conservative: 12
Best Local Similarity: 50.00% Mismatches: 14
Query Match: 48.30% Indels: 0
DB: 1 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-484-710-5 (1-1045)

QY 4 G|nVal|l|eG|yMeTyRAsP|TyRThAlaG|nAsnAsPAsG|uLeuAlaPheAsnLys 23
Db 118 GAGGCCATAGTGAAGTTGACTACCAAGGCCAGACAGATGAGCTGCGATCAAGCGTG 177
QY 24 G|yG|nVal|e|l|eAsnVal|leuAsnLysG|uAsPProAsP|TyR|TyR|ySg|yG|uVal|aSn 43
Db 178 GGTGAATCATCAACCAACATCAGGAAGAGGATGAGGCTGTGGAGGAGACAGATCAAC 237
QY 44 G|yG|nValG|yLeuPheProSeRAsnTyValLys 55
Db 238 GGCAGAGAGGTTTGTTCCTGACCAACTTGTAAAG 273

RESULT 14

US-08-484-709-5
; Sequence 5, Application US/08484709
; Patent No. 5837844
; GENERAL INFORMATION:

APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; ZIP: 02109-1875
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,709
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-192
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1045 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-484-709-5

Alignment Scores:
Pred. No.: 1,11e-13 Length: 1045
Score: 156.00 Matches: 26
Percent Similarity: 73.08% Conservative: 12
Best Local Similarity: 50.00% Mismatches: 14
Query Match: 48.30% Indels: 0
DB: 2 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-08-484-709-5 (1-1045)

QY 4 G|nVal|l|eG|yMeTyRAsP|TyRThAlaG|nAsnAsPAsG|uLeuAlaPheAsnLys 23
Db 118 GAGGCCATAGTGAAGTTGACTACCAAGGCCAGACAGATGAGCTGCGATCAAGCGTG 177
QY 24 G|yG|nVal|e|l|eAsnVal|leuAsnLysG|uAsPProAsP|TyR|TyR|ySg|yG|uVal|aSn 43
Db 178 GGTGAATCATCAACCAACATCAGGAAGAGGATGAGGCTGTGGAGGAGACAGATCAAC 237
QY 44 G|yG|nValG|yLeuPheProSeRAsnTyValLys 55
Db 238 GGCAGAGAGGTTTGTTCCTGACCAACTTGTAAAG 273

RESULT 15

US-08-474-697-5
; Sequence 5, Application US/08474697
; Patent No. 6171800
; GENERAL INFORMATION:
; APPLICANT: Yen-Ming Hsu
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; ZIP: 02109-1875
; COUNTRY: USA
; COMPUTER READABLE FORM:

```

? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/474,697
? FILING DATE:
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Louis Myers
? REGISTRATION NUMBER: 35,965
? REFERENCE/DOCKET NUMBER: BGP-193
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617)227-7400
? TELEFAX: (617)227-5941
? INFORMATION FOR SEQ ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1045 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: both
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? OS-08-474-697-5

```

Alignment Scores:	
Pred. No.:	1,11e-13
Score:	156.00
Percent Similarity:	73.08%
Best Local Similarity:	50.00%
Query Match:	48,308
DB:	3
Length:	1045
Matches:	26
Conservative:	12
Mismatches:	14
Indels:	0
Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) X US-08-474-697-5 (1-1045)

QY	1	GLNValIleelymetyrzsptryrthraIagInanaaPaesGluLeuAlaPheAsnLys	23
Db	118	GAGGCCATAGTGGAGTTTATCCAGAGGCCGACGAGATGAGACTACCATAGCGTGG	177
QY	24	GLYGLNleIleasnValLeubnLysGluAspProAspTrrTyLysGlyValAlaAsn	43
Db	178	GCGGAATCTTCACCAACATCAGAAAGAGAGATGAGAGCGCTGGTGGAGAGACAGATCAAC	23
QY	44	GLYGLNValGlyLeuPheProSerAsnTyValLys	55
Db	238	GGCAGGAGAGAGTTGTTCCTGACCACTTGTGAAGA	273

Search completed: July 1, 2004, 19:54:00
Job time : 45.3138 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:09:20 ; Search time 181.007 Seconds

(without alignments)
1572.471 Million cell updates/sec

Title: US-09-720-934-2_COPY_1080_1138
Perfect score: 323
Sequence: 1 AVOGVIGMDYTRQNDDELA.....GEVNGVGLEPRNYVLTLD 59

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3163042 seqs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool.p/US09720934/funat_30062004_064541_13481/app_query.fasta_1.1386
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=trpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS-human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=Pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09720934_@CGN_1_1500 @runat_30062004_064541_13481
-NCPU=6 -ICPU=3 -NO MMAP -LARGEDUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOP -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database: Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10C_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
------------	-------	-------	-------	--------	----	-------------

1	323	100.0	2067	16	US-10-264-049-887	Sequence 887, App
2	266	82.4	747	9	US-09-879-957-39	Sequence 39, Appl
3	266	82.4	2873	9	US-09-879-957-193	Sequence 193, App
4	266	82.4	3746	11	US-09-764-875-176	Sequence 176, App
5	266	82.4	4053	13	US-10-342-887-1882	Sequence 1882, Ap
6	266	82.4	4053	13	US-10-172-118-1882	Sequence 1882, Ap
7	266	82.4	4210	9	US-09-764-868-125	Sequence 125, Ap
8	266	82.4	5828	13	US-10-398-885A-15	Sequence 15, Appl
9	160	49.5	386	16	US-10-085-783A-37196	Sequence 37196, A
10	160	49.5	400	13	US-10-242-535A-37196	Sequence 37196, A
11	160	49.5	400	13	US-10-085-783A-22639	Sequence 22639, A
12	160	49.5	400	16	US-10-242-535A-22639	Sequence 22639, A
13	160	49.5	1215	17	US-10-648-593-61	Sequence 61, Appl
14	160	49.5	3143	14	US-10-144-621-1	Sequence 1, Appl1
15	160	49.5	3348	13	US-10-342-887-187	Sequence 187, App
16	160	49.5	3348	13	US-10-172-118-187	Sequence 187, App
17	159.5	49.4	2055	16	US-10-085-117-360	Sequence 360, App
18	159.5	49.4	4341	16	US-10-085-117-359	Sequence 359, App
19	158.5	49.1	7696	16	US-10-264-237-2837	Sequence 2837, Ap
20	158.5	49.1	325348	16	US-10-085-117-358	Sequence 358, App
21	158	48.9	184	13	US-10-085-783A-26214	Sequence 26214, A
22	158	48.9	184	16	US-10-242-535A-26214	Sequence 26214, A
23	158	48.9	2967	15	US-10-207-655-197	Sequence 197, App
24	155	48.0	501	9	US-09-764-868-122	Sequence 122, App
25	155	48.0	501	11	US-09-764-875-136	Sequence 136, App
26	155	48.0	501	11	US-09-764-875-444	Sequence 444, App
27	154.5	47.8	2070	16	US-10-085-117-357	Sequence 357, App
28	154.5	47.8	2841	16	US-10-085-117-356	Sequence 356, App
29	154	47.7	1803	9	US-09-962-832-6	Sequence 6, Appl1
30	154	47.7	1803	15	US-10-171-581-43	Sequence 43, Appl
31	153.5	47.5	194945	16	US-10-085-117-355	Sequence 355, App
32	152	47.1	548	10	US-09-918-995-28621	Sequence 28621, A
33	152	47.1	813	17	US-10-641-643-1238	Sequence 1238, Ap
34	152	47.1	984	10	US-09-765-984-5	Sequence 5, Appl1
35	152	47.1	1109	10	US-09-960-706-713	Sequence 713, App
36	152	47.1	1109	15	US-10-007-926A-232	Sequence 232, App
37	152	47.1	1109	15	US-10-327-509-1	Sequence 1, Appl1
38	150	46.4	1661	15	US-10-187-148-2	Sequence 2, Appl1
39	148	45.8	1710	9	US-09-879-957-189	Sequence 189, App
40	146	45.2	476	10	US-09-918-995-35422	Sequence 35422, A
41	146	45.2	500	13	US-10-085-783A-17169	Sequence 17169, A
42	146	45.2	500	16	US-10-242-535A-17169	Sequence 17169, A
43	146	45.2	1889	9	US-09-925-301-274	Sequence 274, App
44	145	44.9	182	15	US-10-029-386-26602	Sequence 26602, A
45	145	44.9	550	15	US-10-029-386-12902	Sequence 12902, A

ALIGNMENTS

RESULT 1
US-10-264-049-887
Sequence 887, Appl
Publication No. US20040005579A1
GENERAL INFORMATION:
APPLICANT: Biase et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA133P1
CURRENT APPLICATION NUMBER: US/10/264,049
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/18569
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: US 60/209,467
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 4360
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 887
LENGTH: 2067
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (5)..(5)

OTHER INFORMATION: n equals a,t,b, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2058)..(2058)
OTHER INFORMATION: n equals a,t,b, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2063)..(2063)
OTHER INFORMATION: n equals a,t,b, or c
US-10-264-049-887

Alignment Scores:

Pred. No.: 4,1e-41 Length: 2067
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-10-264-049-887 (1-2067)

QY 1 AAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
DB 299 GCAGTGTCCAGCGATGGATGGATGACACTACACCGCGCAGAAATGACGATGAGCTGGCC 358
QY 21 PheAsnIysGlyGlnIleIleAsnValLeuAsnIysGluAspProAspTyrTrpIysGly 40
DB 359 TTCAACAGAGGCGCATCATCAACCTCTCAACAGAGGAGCCCTGACTGGTGAAAGGA 418
QY 41 GluValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValIysLeuThrThrasp 59
DB 419 GAATGATGACAGCAAGTGGGCTCTCCCATCCATTTATGTGAAGCTACACAGAC 475

RESULT 2

US-09-879-957-39
Sequence 39, Application US/09879957
Patent No. US20020034755A1
GENERAL INFORMATION:

APPLICANT: SPARKS, Andrew B.
HOFFMAN, No. US20020034755A1h

KAY, Brian K.

FOWLES, Dana M.

MCCONNELL, Stephen J.

TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
USING SAME

NUMBER OF SEQUENCES: 227

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/879,957

FILING DATE: 13-Jun-2001

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/630,915

FILING DATE: 03-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Mistrock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-174

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 747 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-09-879-957-39

Alignment Scores:

Pred. No.: 1.37e-32 Length: 747
Score: 266.00 Matches: 44
Percent Similarity: 91.38% Conservative: 5
Best Local Similarity: 75.86% Mismatches: 0
Query Match: 82.35% Indels: 0
DB: Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-09-879-957-39 (1-747)

QY 2 ValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPhe 21
DB 553 GTATGTCAGTGTGATGCTATGTATGACTATGACCAATATGAAGATGAGCTCGTTTC 612
QY 22 AsnIysGlyGlnIleIleAsnValLeuAsnIysGluAspProAspTyrTrpIysGlyGlu 41
DB 613 TTCAAGGAGCAACCTCATTAATGTTATGACCAAGATGATCTGATTTGGTGAGAGAG 672
QY 42 ValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValIysLeuThrThrasp 59
DB 673 ATCAACGGCGTACGCTGCTCTTCTTCTTCAACACTACGTTAAGATGACACAGAC 726

RESULT 3

US-09-879-957-193
Sequence 193, Application US/09879957
Patent No. US20020034755A1
GENERAL INFORMATION:

APPLICANT: SPARKS, Andrew B.
HOFFMAN, No. US20020034755A1h

KAY, Brian K.

FOWLES, Dana M.

MCCONNELL, Stephen J.

TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
USING SAME

NUMBER OF SEQUENCES: 227

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/879,957

FILING DATE: 13-Jun-2001

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/630,915

FILING DATE: 03-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Mistrock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-174

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

```

;
; TELEEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 193:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 2873 bases
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: unknown
;   MOLECULE TYPE: DNA
;   SEQUENCE DESCRIPTION: SEQ ID NO: 193:
US-09-879-957-193

Alignment Scores:
Pred. No.:      8,45e-32      Length:      2873
Score:          266.00      Matches:      44
Percent Similarity: 91.38%      Conservative: 9
Best Local Similarity: 75.86%      Mismatches: 5
Query Match:      82.35%      Indels:      0
DB:               9          Gaps:      0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-09-879-957-193 (1-2873)

QY      2 ValCyGgInVal1lleglMetYrAspTYrThAlaGlnAsnAspGluLeuAlaPhe 21
        |||
Db      1338 GATATGTCAGGTGATTGCTATGATGACTATGCAGCAATATATGATGATGATGCTCAGTTTC 1397
        |||

QY      22 AsnlyGg1yGln1le1leAsnVal1leuAsnlyGlnAspProAspTyrPlysglyGlu 41
        :
Db      1398 TCCAAAGGACACTCATTAATGTTATGACAAAGATATCTGATTTGGTGGCAAGAGAG 1457
        :

QY      42 ValAsnGlyGlnVal1g1leuPheProSerAsnTYrVal1ysLeuThrThraap 59
        :
Db      1458 ATCAACGGGGTGACTGCTCTTCTTCTTCAACTACGTTAAGATGACGACAGAC 1511
        :

RESULT 4
US-09-764-875-176
; Sequence 176, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
;   APPLICANT: Rosen et al.
;   TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;   FILE REFERENCE: PJ202
;   CURRENT APPLICATION NUMBER: US/09/764,875
;   CURRENT FILING DATE: 2001-01-17
;   Prior application data removed - consult PALM or file wrapper
;   NUMBER OF SEQ ID NOS: 1249
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 176
;   LENGTH: 3746
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-875-176

Alignment Scores:
Pred. No.:      1.21e-31      Length:      3746
Score:          266.00      Matches:      44
Percent Similarity: 91.38%      Conservative: 9
Best Local Similarity: 75.86%      Mismatches: 5
Query Match:      82.35%      Indels:      0
DB:               11         Gaps:      0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-09-764-875-176 (1-3746)

QY      2 ValCyGgInVal1lleglMetYrAspTYrThAlaGlnAsnAspGluLeuAlaPhe 21
        |||
Db      1264 GATATGTCAGGTGATTGCTATGATGACTATGCAGCAATATATGATGATGCTCAGTTTC 1323
        |||

QY      22 AsnlyGg1yGln1le1leAsnVal1leuAsnlyGlnAspProAspTyrPlysglyGlu 41
        :
Db      1324 TCCAAAGGACACTCATTAATGTTATGACAAAGATATCTGATTTGGTGGCAAGAGAG 1383
        :

QY      42 ValAsnGlyGlnVal1g1leuPheProSerAsnTYrVal1ysLeuThrThraap 59
        :
Db      1384 ATCAACGGGGTGACTGCTCTTCTTCTTCAACTACGTTAAGATGACGACAGAC 1437
        :
```

```

RESULT 5
US-10-342-887-1882
; Sequence 1882, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
;   APPLICANT: Dai, Hongyue
;   APPLICANT: He, Yudong
;   APPLICANT: Linsley, Peter S.
;   APPLICANT: Mao, Mao
;   APPLICANT: Roberts, Christopher J.
;   APPLICANT: Van 't Veer, Laura Johanna
;   APPLICANT: Van de Vijver, Marc J.
;   APPLICANT: Bernards, Rene
;   TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;   FILE REFERENCE: 9301-188-999
;   CURRENT APPLICATION NUMBER: US/10/342,887
;   CURRENT FILING DATE: 2003-01-15
;   PRIOR APPLICATION NUMBER: 60/298,918
;   PRIOR FILING DATE: 2001-06-18
;   PRIOR APPLICATION NUMBER: 60/380,710
;   PRIOR FILING DATE: 2002-05-14
;   PRIOR APPLICATION NUMBER: 10/172,118
;   PRIOR FILING DATE: 2002-06-14
;   NUMBER OF SEQ ID NOS: 2639
;   SEQ ID NO 1882
;   LENGTH: 4053
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-342-887-1882

Alignment Scores:
Pred. No.:      1.35e-31      Length:      4053
Score:          266.00      Matches:      44
Percent Similarity: 91.38%      Conservative: 9
Best Local Similarity: 75.86%      Mismatches: 5
Query Match:      82.35%      Indels:      0
DB:               13         Gaps:      0

US-09-720-934-2_COPY_1080_1138 (1-59) x US-10-342-887-1882 (1-4053)

QY      2 ValCyGgInVal1lleglMetYrAspTYrThAlaGlnAsnAspGluLeuAlaPhe 21
        |||
Db      553 GATATGTCAGGTGATTGCTATGATGACTATGCAGCAATATATGATGATGCTCAGTTTC 612
        |||

QY      22 AsnlyGg1yGln1le1leAsnVal1leuAsnlyGlnAspProAspTyrPlysglyGlu 41
        :
Db      613 TCCAAAGGACACTCATTAATGTTATGACAAAGATATCTGATTTGGTGGCAAGAGAG 672
        :

QY      42 ValAsnGlyGlnVal1g1leuPheProSerAsnTYrVal1ysLeuThrThraap 59
        :
Db      673 ATCAACGGGGTGACTGCTCTTCTTCTTCAACTACGTTAAGATGACGACAGAC 726
        :

RESULT 6
US-10-172-118-1882
; Sequence 1882, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
;   APPLICANT: Dai, Hongyue
;   APPLICANT: He, Yudong
;   APPLICANT: Linsley, Peter
;   APPLICANT: Mao, Mao
;   APPLICANT: Roberts, Chris
;   APPLICANT: Van 't Veer, Laura
;   APPLICANT: Van de Vijver, Marc
;   APPLICANT: Bernards, Rene
;   TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;   FILE REFERENCE: 9301-175-999
;   CURRENT APPLICATION NUMBER: US/10/172,118
;   CURRENT FILING DATE: 2002-06-14
;   PRIOR APPLICATION NUMBER: 60/380,770
;   PRIOR FILING DATE: 2002-05-14
;   NUMBER OF SEQ ID NOS: 2699
```

```

; SEQ ID NO 1882
; LENGTH: 4053
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: U61167
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1882

Alignment Scores:
Pred. No.: 1,35e-31
Score: 266.00
Percent Similarity: 91.38%
Best Local Similarity: 75.86%
Query Match: 82.35%
DB: 13

US-09-720-934-2_COPY_1080_1138 (1-59) x US-10-172-118-1882 (1-4053)

QY 2 ValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPhe 21
Db 553 GATATGTCAGGATGATTCCTATGATGACTATGACGCAAAATATGAAGATGAGCTCAGTTTC 612

QY 22 AsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTrpLysGlyGlu 41
Db 613 TCCAAGGACACATCATTAATGATATGAACAAAGATGATCTGATGTTGGCAAGAGAG 672

QY 42 ValAsnGlyGlnValAlGlyLeuPheProSerAsnTyrValLysLeuThrThraSp 59
Db 673 ATCAACGGGGTGACTGCTCTCTCTTCTTCAAACTAGTTAAGATGACGACAGAC 726

RESULT 7
US-09-764-868-125
; Sequence 125, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT732
; CURRENT APPLICATION NUMBER: US/09/764,868
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed -
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 125
; LENGTH: 4210
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-125

Alignment Scores:
Pred. No.: 1.42e-31
Score: 266.00
Percent Similarity: 91.38%
Best Local Similarity: 75.86%
Query Match: 82.35%
DB: 9

US-09-720-934-2_COPY_1080_1138 (1-59) x US-09-764-868-125 (1-4210)

QY 2 ValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPhe 21
Db 1715 GATATGTCAGGATGATTCCTATGATGACTATGACGCAAAATATGAAGATGAGCTCAGTTTC 1774

QY 22 AsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTrpLysGlyGlu 41
Db 1775 TCCAAGGACACATCATTAATGATATGAACAAAGATGATCTGATGTTGGCAAGAGAG 1834

QY 42 ValAsnGlyGlnValAlGlyLeuPheProSerAsnTyrValLysLeuThrThraSp 59
Db 1835 ATCAACGGGGTGACTGCTCTCTCTTCTTCAAACTAGTTAAGATGACGACAGAC 1888

RESULT 8
```

```

US-10-398-885A-15
; Sequence 15, Application US/10398885A
; Publication No. US20040053282A1
; GENERAL INFORMATION:
; APPLICANT: Sugita, Yuji
; APPLICANT: Hashida, Ryochi
; APPLICANT: Ogawa, Kaoru
; APPLICANT: Nagasu, Takeshi
; APPLICANT: Obayashi, Masaya
; APPLICANT: Saito, Hirohisa
; APPLICANT: Takahashi, Eiki
; TITLE OF INVENTION: Method of Testing For Allergic Diseases
; FILE REFERENCE: SHIMIZU-07907
; CURRENT APPLICATION NUMBER: US/10/398,885A
; PRIORITY FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: PCT/JP01/08937
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: JP 2000-314093
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 5828
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: CDS
; LOCATION: (7)..(5052)
; OTHER INFORMATION:
US-10-398-885A-15

Alignment Scores:
Pred. No.: 2.2e-31
Score: 266.00
Percent Similarity: 91.38%
Best Local Similarity: 75.86%
Query Match: 82.35%
DB: 13

US-09-720-934-2_COPY_1080_1138 (1-59) x US-10-398-885A-15 (1-5828)

QY 2 ValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPhe 21
Db 3343 GATATGTCAGGATGATTCCTATGATGACTATGACGCAAAATATGAAGATGAGCTCAGTTTC 3402

QY 22 AsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTrpLysGlyGlu 41
Db 3403 TCCAAGGACACATCATTAATGATATGAACAAAGATGATCTGATGTTGGCAAGAGAG 3462

QY 42 ValAsnGlyGlnValAlGlyLeuPheProSerAsnTyrValLysLeuThrThraSp 59
Db 3463 ATCAACGGGGTGACTGCTCTCTCTTCTTCAAACTAGTTAAGATGACGACAGAC 3516

RESULT 9
US-10-085-783A-37196
; Sequence 37196, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; PRIORITY FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 13:01:26 ; Search time 1774.84 Seconds

(Without alignments)
992.694 Million cell updates/sec

Title: US-09-720-934-2_COPY_1080_1138
Perfect score: 323
Sequence: 1 AVCGVIGMYDYTAQNDELAA.....GEVNGVGLPSPNYKLTDD 59

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_P/US09720934/runat_30062004_064540_13442/app_query.fasta_1.1386
-DB=EST-QFMT=fastap -SUFFIX=rst -MIMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCAIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEADSIZE=500 -MINLEN=2000000000
-USER=US09720934.@CNC_1_1.12431.@runat_30062004_064540_13442 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST.*
1: em_estba.*
2: em_esthum.*
3: em_estlm.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_luv.*
19: em_gss_pin.*
20: em_gss_vit.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rdg.*
26: em_gss_pig.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	323	100.0	487	9	AI151142
2	323	100.0	545	14	CB161912
3	323	100.0	575	14	CB163763
4	323	100.0	583	14	CB156803
5	323	100.0	591	14	CB215044
6	323	100.0	1062	13	BO898718
7	323	100.0	1064	10	BE888832
8	319	98.8	363	14	H66900
9	317	98.1	581	12	BG894858
10	317	98.1	612	12	BG894850
11	315	97.5	693	13	BU321701
12	313	96.9	623	13	BU056026
13	313	96.9	634	13	BU054334
14	313	96.9	782	12	BG922829
15	313	96.9	782	13	BU614534
16	313	96.9	906	13	BU519029
17	313	96.9	2589	11	AK082606
18	313	96.9	5385	11	BC062938
19	312	96.6	538	14	CA390605
20	312	96.6	1073	13	BQ432912
21	308	95.4	406	9	AV590688
22	308	95.4	594	14	CB172533
23	308	95.4	612	14	CB435542
24	308	95.4	629	14	CF768246
25	306	94.7	603	9	AI594919
26	301.5	93.3	521	12	BG876696
27	295	91.3	591	12	BU639774
28	295	91.3	713	14	CB209507
29	295	91.3	898	14	CA987804
30	293	90.7	749	13	BU963831
31	293	90.7	830	14	CB988722
32	283	87.6	674	13	BU054591
33	281	87.0	539	9	AL915730
34	276	85.4	1036	29	CNS02AR3
35	272	84.2	552	12	BM247316
36	268	83.0	684	14	CA365705
37	266	82.4	225	14	CF135884
38	266	82.4	372	10	BP918427
39	266	82.4	455	12	BM757984
40	266	82.4	458	12	BI031994
41	266	82.4	517	12	BM750308
42	266	82.4	640	12	BM758638
43	266	82.4	702	13	BX502192
44	265	82.0	549	12	BM233591
45	265	82.0	576	10	AM631635

ALIGNMENTS

RESULT 1
AI151142/c
LOCUS AI151142 487 bp mRNA linear EST 28-OCT-1998
DEFINITION qc87e04.x1 Soares pregnant uterus NbHPV Homo sapiens
IMAGE:1721214 3' similar to SW:SH17 HUMAN Q15811 SH3
DOMAIN-CONTAINING PROTEIN SH3p17. ; mRNA sequence.
ACCESSION AI151142
VERSION AI151142.1 GI:3679611
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 487)
 AUTHORS NC1-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 518 Std Error: 0.00
 Seq primer: -40m13 fwd. fr from AmerSham
 High quality sequence stop: 444.
 Location/Qualifiers
 source
 1..487
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:1721214"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_idb="Scotres_pregnant_uterus_NBHPV"
 /note="Organ: uterus; Vector: pT7T3-Pac; Site 1: Not I;
 Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer [5',
 AACTGGAAGATTGCGCGCCCTTTTCTTTTCTTTT 3']
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization. Library
 constructed by M. Fatima Bonaldo."

ORIGIN
 Alignment Scores:
 Pred. No.: 2,83e-36 Length: 487
 Score: 323.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x A1151142 (1-487)

QY 1 A1A1A1CysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
 Db 430 GCAGTGTCCAGGCGATGGATGTACACTACACCGCGAGATGACGATGAGCTGGCC 371

QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTrpLysGly 40
 Db 370 TTCAACAAGGCGCCAGATCATCACTCTCTCAACAAGGAGGACCCGACTGGTGGAAAGA 311

QY 41 GluValaAnglyGlnValaGlyLeuPheProSerAsnTyrValLysLeuThrThrasp 59
 Db 310 GAAGTCATGACAGATGGGGCTCTTCCCATCCCAATTATGTGAAGCTGACACACAGAC 254

RESULT 2
 LOCUS CB161912 545 bp mRNA linear EST 30-JAN-2003
 DEFINITION K-EST0221969 L17N670205n1 Homo sapiens cDNA clone
 L17N670205n1-41-H01 5', mRNA sequence.
 ACCESSION CB161912
 VERSION CB161912.1 GI:28148038
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 545)
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.U., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)

COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 41 row: H column: 01
 High quality sequence stop: 545.
 Location/Qualifiers
 source
 1..545
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_image="L17N670205n1-41-H01"
 /sex="F"
 /lab_host="TOP10F"
 /clone_idb="L17N670205n1"
 /note="Organ: Liver; Vector: pT7T3-Pac; Site 1: EcoRI;
 Site 2: NotI; The library was contributed by the Soares
 laboratory and it was constructed as described by Bonaldo,
 M.F., Lemmon, G. and Soares, M.B. (1996). Genome Research
 6(9): 791-806. RNA was prepared from harvested cell
 culture."

ORIGIN
 Alignment Scores:
 Pred. No.: 3.34e-36 Length: 545
 Score: 323.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x CB161912 (1-545)

QY 1 A1A1A1CysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
 Db 361 GCAGTGTCCAGGCGATGGATGTACACTACACCGCGAGATGACGATGAGCTGGCC 420

QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTrpLysGly 40
 Db 421 TTCAACAAGGCGCCAGATCATCACTCTCTCAACAAGGAGGACCCGACTGGTGGAAAGA 480

QY 41 GluValaAnglyGlnValaGlyLeuPheProSerAsnTyrValLysLeuThrThrasp 59
 Db 481 GAAGTCATGACAGATGGGGCTCTTCCCATCCCAATTATGTGAAGCTGACACACAGAC 537

RESULT 3
 LOCUS CB163763 575 bp mRNA linear EST 30-JAN-2003
 DEFINITION K-EST0224708 L17N670205n1 Homo sapiens cDNA clone
 L17N670205n1-46-D04 5', mRNA sequence.
 ACCESSION CB163763
 VERSION CB163763.1 GI:28149889
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 575)
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.U., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr

Plate: 46 row: D column: 04
High quality sequence stop: 575.
Location/Qualifiers
1..575

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="U17N670205n1-46-D04"
/sex="F"
/lab_host="Top10F"
/clone_id="U17N670205n1"
/note="Organ: Liver; Vector: pT73-Pac; Site 1: EcORI;
Site 2: NotI; The library was contributed by the Soares
Laboratory and it was constructed as described by Bonaldo,
M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
6(9): 791-806. RNA was prepared from harvested cell
culture."

ORIGIN

Alignment Scores:

Pred. No.: 3,61e-36 Length: 575
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x CB156803 (1-575)

QY 1 AlValaCysGlnValIleGlyMetTyrAspTyrThrAglInsAspAspGluLeuAla 20
DB 361 GCAGTGTGCCAGGTGATGGATGTACGACTACCGCGGAGAAATGACGATGAGCTGGCC 420
QY 21 PheAsnLysGlyGlnIlelleAsnValLeuAsnLysGluAspProAspTyrTyrPlyGly 40
DB 421 TTCACACAGGGCCAGATCATCAACGTCCTCACACAGAGAGACCCCTGACTGGTGAAGA 480
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59
DB 481 GAAGTCAATGACAAAGTGGGGCTCTTCCCATCAATTATGTGAAGTGACACAGAC 537

RESULT 4 CB156803 583 bp mRNA linear EST 29-JAN-2003
LOCUS K-EST0215771 U17N670205n1 Homo sapiens cDNA clone
DEFINITION U17N670205n1-4-C07 5', mRNA sequence.
ACCESSION CB156803
VERSION CB156803.1 GI:28141931
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 583) Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 4 row: C column: 07
High quality sequence stop: 583.

FEATURES

source 1..583
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="U17N670205n1-4-C07"
/sex="F"
/lab_host="Top10F"
/clone_id="U17N670205n1"
/note="Organ: Liver; Vector: pT73-Pac; Site 1: EcORI;
Site 2: NotI; The library was contributed by the Soares
Laboratory and it was constructed as described by Bonaldo,
M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
6(9): 791-806. RNA was prepared from harvested cell
culture."

ORIGIN

Alignment Scores:

Pred. No.: 3,68e-36 Length: 583
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x CB156803 (1-583)

QY 1 AlValaCysGlnValIleGlyMetTyrAspTyrThrAglInsAspAspGluLeuAla 20
DB 361 GCAGTGTGCCAGGTGATGGATGTACGACTACCGCGGAGAAATGACGATGAGCTGGCC 420
QY 21 PheAsnLysGlyGlnIlelleAsnValLeuAsnLysGluAspProAspTyrTyrPlyGly 40
DB 421 TTCACACAGGGCCAGATCATCAACGTCCTCACACAGAGAGACCCCTGACTGGTGAAGA 480
QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59
DB 481 GAAGTCAATGACAAAGTGGGGCTCTTCCCATCAATTATGTGAAGTGACACAGAC 537

RESULT 5 CB215044 591 bp mRNA linear EST 06-FEB-2003
LOCUS NISC_np01906.y1 NICHD_HS_U1 Homo sapiens cDNA clone IMAGE:5936818
DEFINITION 5', mRNA sequence.
ACCESSION CB215044
VERSION CB215044.1 GI:28263236
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 591) Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
CDNA Library Preparation:
DNA Sequencing By: The I.M.A.G.E. Consortium/LLNL
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Plate: LLNL13162 row: M column: 11
Seq primer: M13RP reverse primer (ABI).
Location/Qualifiers
1..591

FEATURES

source 1..591
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5936818"
/sex="female"
/tissue_type="normal endometrium, late proliferative
phase, cycle day 13"
/lab_host="DH10B (T1-resistant)"
/clone_id="NICHD_HS_U1"

/note="Organ: uterus; Vector: PCMV-SPORT6.1.ccdh (ResGen, Invitrogen Corporation); Site 1: NotI; Site 2: EcoRV; Cloned unidirectionally from microquantity amounts of mRNA from normal endometrial tissue (late proliferative phase, cycle day 13). Average insert size 1.9 kb. Library constructed by ResGen (Invitrogen Corporation)."

ORIGIN

Alignment Scores:
Pred. No.: 3,75e-36 Length: 591
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x CB215044 (1-591)

QY 1 ALaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20

Db 153 GCACTGTGCCAGGCGATTGGGATTTACGACTACCCGCGCAGATGACGATGAGCTGGCC 212

QY 21 PheAsnlysglyGlnIleIleAsnValLeuAsnlysgluAspProAspTyrTrpLysGly 40

Db 213 TTCAACAAGGCGCCAGATCATCAACGCTCTCAACAAGAGGACCTGACTGATGGAAGA 272

QY 41 GluValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValIysLeuThrThrAsp 59

Db 273 GAAGTCATATGACACAGTGGGGCTCTTCCATCCAAATTATGTGAAGCTGACACAGAC 329

RESULT 6

LOCUS BQ898718 1062 bp mRNA linear EST 16-AUG-2002

DEFINITION AGENCOURT 8118759 lupski dorsal root ganglion Homo sapiens cDNA clone IMAGE:6180303 5', mRNA sequence.

ACCESSION BQ898718

VERSION BQ898718.1 GI:22290732

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens (human); Chordata; Craniata; Vertebrata; Euteleostomi;

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1062)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cga@bbs-remail.nih.gov

Tissue Procurement: Dr. James R. Lupski

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov

Plate: L1AM13562 row: n column: 16

High quality sequence start: 116

High quality sequence stop: 760.

Location/Qualifiers

1..1062

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6180303"

/sex="male"

/tissue_type="dorsal root ganglia"

/dev stage="adult, 36 yr"

/lab host="DH10B"

/clone_lib="lupski dorsal root ganglion"

/note="Vector: PCMV-SPORT6 (Life Technologies); Site 1:

NotI; Site 2: SalI; cDNA made by oligo-dT priming.

Directionally cloned using the following adaptors:

5'-TCGACCCACGCGTCCG-3' and

5'-GACTAGTTCAGATCCGACGCCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

ORIGIN

Alignment Scores:
Pred. No.: 8,82e-36 Length: 1062
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x BQ898718 (1-1062)

QY 1 ALaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20

Db 574 GCACTGTGCCAGGCGATTGGGATTTACGACTACCCGCGCAGATGACGATGAGCTGGCC 633

QY 21 PheAsnlysglyGlnIleIleAsnValLeuAsnlysgluAspProAspTyrTrpLysGly 40

Db 634 TTCAACAAGGCGCCAGATCATCAACGCTCTCAACAAGAGGACCTGACTGATGGAAGA 693

QY 41 GluValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValIysLeuThrThrAsp 59

Db 694 GAAGTCATATGACACAGTGGGGCTCTTCCATCCAAATTATGTGAAGCTGACACAGAC 750

RESULT 7

LOCUS BE888832 1064 bp mRNA linear EST 20-OCT-2000

DEFINITION 60151375P1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3914940 5', mRNA sequence.

ACCESSION BE888832

VERSION BE888832.1 GI:10345530

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens (human); Chordata; Craniata; Vertebrata; Euteleostomi;

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1064)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cga@bbs-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov

Plate: L1AM9737 row: h column: 13

High quality sequence stop: 607.

Location/Qualifiers

1..1064

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3914940"

/tissue_type="leiomyosarcoma"

/lab host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_71"

/note="Organ: uterus; Vector: PCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 2.1 kb.

ORIGIN

Alignment Scores: 8,85e-36 Length: 1064
Pred. No.: 8,85e-36 Length: 1064

Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x BB888832 (1-1064)

QY 1 A1aValCysGlnVal111eglyMetTyrAspTyrThrAlaGlnAsnAspGluLeuAla 20
Db 122 GCAGTGGCCAGGAGGATGGATGATGACACTACACCCGCGAGATGACGATGAGCTGGCC 181
QY 21 PheAsnLysGlyGlnIle11eAsnValLeuAsnLysGluAspProAspTyrTyrLysGly 40
Db 182 TTCACAGAGGCCAGATCATCAAGCTCTCTCAACAGAGAGGAGCCTGACTGGTGAAGA 241
QY 41 GluValAsnGlyGlnVal1GlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59
Db 242 GAAGTCAATGAGACAGAGGGGCTCTTCCCATCCCATTTATGTGAAGCTGACACAGAC 298

RESULT 8
LOCUS H66900 363 bp mRNA linear EST 18-OCT-1995
DEFINITION y71d07.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:210733 5' similar to SP:MYSC_ACACA P10569 MYOSIN IC HEAVY ;,
H66900
H66900.1 GI:1025640
EST.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM

ACCESSION H66900
VERSION H66900.1
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estw@wustl.edu
Insert Size: 522
High quality sequence stops: 299
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert length: 522 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 299.
Location/Qualifiers
1. 363
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3783574"
/db_xref="taxon:9606"
/clone="IMAGE:210733"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer

[5' AACTGGAGAAATTATTAAGATCTTTTCTTTTCTTTT 3']
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Alignment Scores:
Pred. No.: 6,97e-36 Length: 363
Score: 319.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.76% Indels: 0
DB: 14 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x H66900 (1-363)

QY 2 ValCysGlnVal111eglyMetTyrAspTyrThrAlaGlnAsnAspGluLeuAlaPhe 21
Db 108 GTGTGCCAGGTGATGGATGATGACACTACACCCGCGAGATGACGATGAGCTGGCTTC 167
QY 22 AsnLysGlyGlnIle11eAsnValLeuAsnLysGluAspProAspTyrTyrLysGlyGlu 41
Db 168 AACAGAGGCCAGATCATCAAGCTCTCTCAACAGAGAGCCTGACTGGTGAAGA 227
QY 42 ValAsnGlyGlnVal1GlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59
Db 228 GTCAATGAGACAGAGGGGCTCTTCCCATCCCATTTATGTGAAGCTGACACAGAC 281

RESULT 9
LOCUS BG894858 581 bp mRNA linear EST 05-JUN-2001
DEFINITION BG894858
ACCESSION BG894858
VERSION BG894858.1 GI:14305099
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCAGTCACGACG
Plate: 120 row: G column: 3
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers
1. 581
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1P1G"
/note="Vector: pcwv SPOR6; Site_1: NotI; Site_2: SalI;

FEATURES
source
1. 581
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1P1G"
/note="Vector: pcwv SPOR6; Site_1: NotI; Site_2: SalI;

Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

ORIGIN

Alignment Scores:
Pred. No.: 2.69e-35 Length: 581
Score: 317.00 Matches: 57
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 96.61% Mismatches: 0
Query Match: 98.14% Indels: 0
DB: 12 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x BG894858 (1-581)

QY 1 AAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
DB 167 GCAGTGTCCACGGCGATCGGATGACACTACACCCGGCAGAAAGACGAGAGCTGGCC 226

QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTriPlyGly 40
DB 227 TTCAACAAGGGCCAGGTATCATCACTCTCAACAAGAGAGACCCGACTGGTGGAGGGA 286

QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrasp 59
DB 287 GAAGTCCACGGCCAAAGTGGGGCTTCCCGTCCAACTACGTGAAGCTGACCAACAGAC 343

RESULT 10
BG894850 612 bp mRNA linear EST 05-JUN-2001
LOCUS 355615 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
DEFINITION BG894850
ACCESSION BG894850.1 GI:14305091
VERSION
KEYWORDS
SOURCE
ORGANISM
Sus scrofa (pig)

REFERENCE
AUTHORS
TITLE
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 612)
Fahrenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,
Vallet, J., Wise, T., Rohrer, G.A., Petrea, G., Sultana, R.,
Quackenbush, J., and Keele, J.W.
Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)

JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -mismatch 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCCACTGACGACG
Plate: 120 row: B column: 4
Seq primer: ATTAGGTGACCTATAG.
Location/Qualifiers

FEATURES
Source
1..612
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1P1G"
/note="Vector: pCMV SPOR6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

ORIGIN

Alignment Scores:

Pred. No.: 2.9e-35 Length: 612
Score: 317.00 Matches: 57
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 96.61% Mismatches: 0
Query Match: 98.14% Indels: 0
DB: 12 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x BG894850 (1-612)

QY 1 AAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
DB 167 GCAGTGTCCACGGCGATCGGATGACACTACACCCGGCAGAAAGACGAGAGCTGGCC 226

QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTriPlyGly 40
DB 227 TTCAACAAGGGCCAGGTATCATCACTCTCAACAAGAGAGACCCGACTGGTGGAGGGA 286

QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrasp 59
DB 287 GAAGTCCACGGCCAAAGTGGGGCTTCCCGTCCAACTACGTGAAGCTGACCAACAGAC 343

RESULT 11
BU321701 693 bp mRNA linear EST 28-NOV-2002
LOCUS 603486370P1 CSEQCHN62 Gallus gallus cDNA clone CHEST381118 5', mRNA
DEFINITION
ACCESSION BU321701.1 GI:25829702
VERSION
KEYWORDS
SOURCE
ORGANISM
Gallus gallus (chicken)

REFERENCE
AUTHORS
TITLE
Phasianinae; Gallus.
1 (bases 1 to 693)
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.D.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)

JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 0161208930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers

FEATURES
Source
1..693
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST381118"
/dev_stage="36"
/lab_host="DH10B"
/clone_lib="CSEQCHN62"
/note="Organ: heads; Vector: pBluescript II KS(+); Site 1:
EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996) 791, except that a significantly longer

ORIGIN reannealing hybridization was used."

Alignment Scores:
 Pred. No.: 6.75e-35 Length: 693
 Score: 315.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 2
 Best Local Similarity: 96.61% Mismatches: 0
 Query Match: 97.52% Indels: 0
 DB: 13 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x BU321701 (1-693)

QY 1 A1AValCysGlnVal11leGlyMetTyrAspTyrThrAlaGlnAspAspGluLeuAla 20
 Db 101 TCAGTGTGCCAAGTAATGGCATGTATGACTACACGACAGATGATGATGAGTACGCA 160
 QY 21 PheAsnLysGlyGlnIle1leAsnVal1leAsnLysGluAspProAspTyrTyrPlyGly 40
 Db 161 TTCAGTAAGGCCACAGATTATAACGTCTTAAACAGAGAGACCCAGACTGTGTGAAGG 220
 QY 41 GluValAsnGlyGlnVal1GlyLeuPheProSerAsnTyrValLysLeuThrAsp 59
 Db 221 GAAGTGAATGGACAAGTGGCTCTTCCATCCACATATGTGAAGCTGACACACAGAT 277

RESULT 12 623 bp mRNA linear EST 26-AUG-2002
 BU056026 UI-M-FPO-bzu-m-16-0-UI.r1 NIH_BMAP_FPO Mus musculus cDNA clone
 LOCUS IMAGE:6407607 5', mRNA sequence.
 BU056026
 VERSION BU056026.1 GI:224946103
 KEYWORDS EST
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/MLN at:
 http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: PYX-5.
 Location/Qualifiers
 1..623

FEATURES
 source
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6407607"
 /tissue_type="whole brain"
 /dev_stage="embryo 15dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_id="NIH_BMAP_FPO"
 /note="Organ: Brain; Vector: PYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 to Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into PYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail
 is GTGCGTGGAA. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

Alignment Scores:
 Pred. No.: 1.12e-34 Length: 623
 Score: 313.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 2
 Best Local Similarity: 96.61% Mismatches: 0
 Query Match: 96.90% Indels: 0
 DB: 13 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x BU056026 (1-623)

QY 1 A1AValCysGlnVal11leGlyMetTyrAspTyrThrAlaGlnAspAspGluLeuAla 20
 Db 47 GCAGTGTGCCAAGTATCGGATGTACGATTACACCGCCAGAACGATGACGACTAGCC 106
 QY 21 PheAsnLysGlyGlnIle1leAsnVal1leAsnLysGluAspProAspTyrTyrPlyGly 40
 Db 107 TTCAGTAAGGCCACAGATTATAACGTCTTAAACAGAGAGACCCAGACTGTGTGAAGG 166
 QY 41 GluValAsnGlyGlnVal1GlyLeuPheProSerAsnTyrValLysLeuThrAsp 59
 Db 167 GAAGTGAATGGACAAGTGGCTCTTCCATCCATATATGTAAAGCTGACACACAGAC 223

RESULT 13 694 bp mRNA linear EST 26-AUG-2002
 BU054334 UI-M-FPO-bzh-d-08-0-UI.r1 NIH_BMAP_FPO Mus musculus cDNA clone
 LOCUS IMAGE:6403783 5', mRNA sequence.
 BU054334
 VERSION BU054334.1 GI:22494411
 KEYWORDS EST
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/MLN at:
 http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: PYX-5.
 Location/Qualifiers
 1..694

FEATURES
 source
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6403783"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5 dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_id="NIH_BMAP_FPO"
 /note="Organ: brain; Vector: PYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TGAGAGAGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): "Gene Discovery in the Developing Mouse Nervous System", supported by National Institutes of Mental Health (NIMH), Hemm Chinn, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
 Pred. No.: 1,316-34
 Score: 313.00
 Percent Similarity: 100.00%
 Best Local Similarity: 96.61%
 Query Match: 96.90%
 DB: 13

US-09-720-934-2_COPY_1080_1138 (1-59) x BU054334 (1-694)

QY 1 AAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspGluLeuAla 20
 Db 357 GCAAGTGTGCCAGGTGATCGGATCGGATGACATTACACCCCGAGAACGATGAGAACTAGCC 416
 QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTyrPlyGly 40
 Db 417 TTTCAGCAAAAGCCAGATCATCATCGTCTCTCAACAAGAGAGACCCGAGCTGTGGAAAGGA 476

QY 41 GluValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59
 Db 477 GAAGTCAAGTGGGCAAGTGGGCTCTTCCATCCCAATTATGTAAGCTGACCAAGAC 533

RESULT 14

LOCUS BG922629 782 bp mRNA linear EST 05-JUN-2001
 DEFINITION 602821583F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4950524 5',
 mRNA sequence.
 ACCESSION BG922629
 VERSION BG922629.1 GI:14303105
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: csagbbs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: InCyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/BLN at:
 http://image.llnl.gov
 Plate: L14M10905 row: 6 column: 21
 High quality sequence shop: 716.

FEATURES

Location/Qualifiers
 1..782
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGR:4950524"
 /sex="female, virgin"
 /tissue_type="infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"

/clone.lib="NCI CGAP Mam6"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI. Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Jeffrey Green, M.D., NIH"

ORIGIN

Alignment Scores:
 Pred. No.: 1,566-34
 Score: 313.00
 Percent Similarity: 100.00%
 Best Local Similarity: 96.61%
 Query Match: 96.90%
 DB: 12

US-09-720-934-2_COPY_1080_1138 (1-59) x BG922629 (1-782)

QY 1 AAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspGluLeuAla 20
 Db 138 GCAAGTGTGCCAGGTGATCGGATCGGATGACATTACACCCCGAGAACGATGAGAACTAGCC 197

QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTyrPlyGly 40
 Db 198 TTTCAGCAAAAGCCAGATCATCATCGTCTCTCAACAAGAGAGACCCGAGCTGTGGAAAGGA 257

QY 41 GluValAsnGlyGlnValIleGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59
 Db 258 GAAGTCAAGTGGGCAAGTGGGCTCTTCCATCCCAATTATGTAAGCTGACCAAGAC 314

RESULT 15

LOCUS BU614534 782 bp mRNA linear EST 20-FEB-2003
 DEFINITION U1-M-EV0-cbg-1-22-0-UI.r1 NIH BMAP EV0 Mus musculus cDNA clone
 U1-M-EV0-cbg-1-22-0-UI 5', mRNA sequence.
 ACCESSION BU614534
 VERSION BU614534.1 GI:23280749
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: csagbbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@iowa.edu
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES

source

Seq primer: pYX-5.
 Location/Qualifiers
 1..782
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="U1-M-EV0-cbg-1-22-0-UI"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone.lib="NIH-BMAP EV0"

/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded

cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pUX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTGGGTGGA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemmi Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.:	1.56e-34	Length:	782
Score:	313.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	2
Best Local Similarity:	96.61%	Mismatches:	0
Query Match:	96.90%	Indels:	0
DB:	13	Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) x BU614534 (1-782)

QY	1	AlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspGluLeuAla	20
Db	427	GCAGTGTGCCAGGTATCGGATGTACGATTACACGCCGACGATGACGAATTAAGCC	486
QY	21	PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrPylLysGly	40
Db	487	TTACAGCAAGGCCAGATCATCAAGTCTCTCAACAAGAGGACCCGACTGGTGGANAGAA	546
QY	41	GluValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp	59
Db	547	GAAGTCAGTGGGCAAGTGGGCTCTCCATCCATTATGTAAAGCTGACCAACAGAC	603

Search completed: July 1, 2004, 19:49:00
Job time : 1780.84 secs

(UYN-) UNITV NORTH CAROLINA.
XX Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;
XX WPI: 1996-465045/46.
XX P-PSDB; AAM05395.
XX Identifying polypeptide(s) having specific functional domain (esp. SH3
XX domain) - comprises detecting selective binding to recognition unit,
XX regardless of sequence homology.
XX Claim 76; Fig 50; 174pp; English.
XX
XX AAT39786-T39803 represent novel human and mouse genes encoding Src-
XX homology region 3 (SH3) domain containing proteins that can be used in
XX the method of the invention. SH3 domain containing proteins play a role
XX in signalling and structural elements of cells. The method of the
XX invention is for identifying polypeptides containing functional domains
XX of interest (especially SH3 domains). The method comprises contacting a
XX multivalent recognition unit (RU) complex with a number of peptides and
XX identifying polypeptides having a selective binding affinity for the RU
XX complex. The method is based on functional similarities and does not rely
XX on sequence similarities. Prior methods only gave limited success for
XX identifying proteins which contain an SH3 domain due to the minimal
XX sequence homology among known SH3 proteins. It has been found that small
XX peptide RUS in multivalent form have reduced specificity for a given
XX functional domain compared to monomer RUS. Multivalent RU complexes are
XX particularly suited to screening for polypeptides containing functional
XX domains that are similar to, but not identical in sequence to, the
XX original target functional domain. The new method enables proteins having
XX a common function to be identified. Identification of novel SH3 proteins
XX will be useful for a better understanding of cell growth, malignancy,
XX signal transduction processes, etc. New candidate drugs can be
XX identified, and their specificities (e.g. pharmacological activities) can
XX be assessed using the method of the invention
XX
XX Sequence 1389 BP; 404 A; 370 C; 357 G; 258 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 6.9e-35 Length: 1389
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-720-934-2_COPY_740_800 (1-61) x AAT39795 (1-1389)
QY 1 ValysValValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIle 20
Db 172 GTAAAGGAGGTATTTACCGGACCTGACCCCTTGATCCAGAGCCATGATGAAATC 231
QY 21 ThrileGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr 40
Db 232 ACTATCCAGCCGAGGACATGATCATGTGATGAAAGCAAACTGAGAAACCCGCTGG 291
QY 41 LeuGlyGlyGluLeuLeuLeuGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIle 60
Db 232 CTTGAGAGGAGATTAAGGAAAGACAGGGTGTTCCCTGCAAACTATGCAAGAAATC 351
QY 61 Pro 61
Db 352 CCA 354
RESULT 2
AAS84762
ID AAS84762 standard; cDNA; 2874 BP.
XX AAS84762;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #20566.

XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Dermanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
XX P-PSDB; ABG20575.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 1; SEQ ID NO 20566; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS6197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 2874 BP; 940 A; 679 C; 703 G; 552 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.7e-34 Length: 2874
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-720-934-2_COPY_740_800 (1-61) x AAS84762 (1-2874)
QY 1 ValysValValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIle 20
Db 913 GTAAAGGAGGTATTTACCGGACCTGACCCCTTGATCCAGAGCCATGATGAAATC 972
QY 21 ThrileGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr 40
Db 973 ACTATCCAGCCGAGGACATGATCATGTGATGAAAGCAAACTGAGAAACCCGCTGG 1032
QY 41 LeuGlyGlyGluLeuLeuLeuGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIle 60

Db 1033 CTTGAGGAGAAATTAAAGAAAGACAGGGTGTCTCCCTGCAGAACTATGACAGAAAATC 1092

QY 61 Pro 61
|||
Db 1093 CCA 1095

RESULT 3
AAZ34574
ID AAZ34574 standard; cDNA; 3231 BP.

XX AC AAZ34574;
XX
DT 01-FEB-2000 (first entry)
XX
DE Human SH3D1A cDNA clone 9.
XX

KM SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
KM megakaryocytic abnormality; myeloproliferative disorder;
KM platelet disorder; neural disorder; thrombocytopenia;
KM haematopoietic disorder; cognitive dysfunction; microcephaly;
KM lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
OS Homo sapiens.
XX

XX Key Location/Qualifiers
FT CDS 2..1927
FT /*tag= a
XX

PN MO9953062-A2.
XX
PD 21-OCT-1999.
XX
PF 16-APR-1999; 99WC-US008371.
XX
PR 16-APR-1998; 98US-0082007P.
XX
PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
XX
PI Korenberg JR, Chen X;
XX
DR MPI: 1999-633829/54.
XX
DR P-PSDB; AAY32158.
XX

PT Nucleic acid from the human SH3D1A gene and its products, useful for the
PT diagnosis and treatment of myeloproliferative disorders and leukemia.
XX
PS Claim 2; Fig 14; 99p; English.
XX

CC This is the nucleotide sequence of a non-full-length cDNA (clone 9)
CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
CC contributes to the development of platelets and the pathogenesis of
CC leukaemias, both in general and in particular those involving the
CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
CC region for low platelets on chromosome 21. Sequencing of 5 different
CC sizes of cDNA clone from foetal brain (see AAZ34570-74) suggests that at
CC least 3 isoforms exist. The invention provides methods for the diagnosis
CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
CC association of gains in chromosome 21, low platelets in deletion for 21,
CC abnormalities, dysfunctions and disorders including brain malformations
CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly
CC and colpocephaly. Methods are also provided for: suppressing cells unable
CC to regulate themselves; screening for a somatic alteration in the SH3D1A
CC gene; monitoring the progress and adequacy of a treatment; monitoring
CC tumour risk progress or megakaryocytic abnormality; myeloproliferative
CC disorder; haematopoietic disorder; platelet disorder or leukaemia; and
CC treatment of a subject (including a prenatal subject) having
CC megakaryocytic abnormality, myeloproliferative disorder, platelet
CC disorder, leukaemia or neural disorder using a nucleic acid that
CC expresses SH3D1A or its antisense nucleic acid
XX
XX Sequence 3231 BP; 1004 A; 721 C; 712 G; 794 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.97e-34 Length: 3231
Score: 333.00
Percent Similarity: 100.00% Matches: 61
Best Local Similarity: 100.00% Conservative: 0
Query Match: 100.00% Mismatches: 0
DB: 2 Indels: 0
Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AAZ34574 (1-3231)

QY 1 ValysValValYTYrTgAlaLeuTYrProPheGluSerArgSerHisAspGluIle 20
Db 710 GTAAAGTGTGATATCCGGGCACTGTACCCCTTGAATCCAGAACCATGATGAATC 769

QY 21 ThrTLeGInProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr 40
Db 770 ACTATCCAGCCAGAGACATGATGCTAGTGGATGTAAGCCAAACTGGAGAAACCGGCTGG 829

QY 41 LeuGlyGlyGluLeuValysGlyTyrThrGlyTyrPheProAlaAenTYrAlaGluIle 60
Db 830 CTTGAGGAGAAATTAAAGAAAGACAGGGTGTCTCCCTGCAGAACTATGACAGAAAATC 889

QY 61 Pro 61
|||
Db 890 CCA 892

RESULT 4
ABK3498
ID ABK3498 standard; cDNA; 3319 BP.
XX
AC ABK3498;
XX
DT 05-JUN-2002 (first entry)
XX
DE DNA encoding novel central nervous system protein #78.
XX

KM Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KM hyperproliferative disorder; neoplasia; cardiovascular disorder;
KM cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KM nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KM acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KM adenocarcinoma; reproductive system disorder; testicular feminisation;
KM endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KM respiratory disorder; renal disorder; kidney failure; blood disorder;
KM myocardial infarction; wound healing; cell proliferation; skin aging;
KM food additive; food preservative; gene therapy; gene; ss.
XX
XX Homo sapiens.
XX
XX WO20015318-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WC-US001332.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
XX 02-MAR-2000; 2000US-0186350P.
XX
XX 16-MAR-2000; 2000US-0189874P.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX
XX 18-APR-2000; 2000US-0198123P.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX 28-JUN-2000; 2000US-0214886P.
XX
XX 30-JUN-2000; 2000US-0215135P.
XX
XX 07-JUL-2000; 2000US-0216647P.
XX
XX 07-JUL-2000; 2000US-0216880P.
XX
XX 11-JUL-2000; 2000US-0217487P.
XX
XX 11-JUL-2000; 2000US-0217496P.
XX
XX 14-JUL-2000; 2000US-0218290P.
XX
XX 26-JUL-2000; 2000US-0220963P.

CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infection. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,

Alignment Scores:

Pred. NO.:	2.04e-34	Length:	3319
Score:	333.00	Matches:	61
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-720-934-2_COPY_740_800 (1-61) x ABR43498 (1-3319)

QY 1 VallystValValYTYTYRARGAlaLeuTYRProPheGluSerArgSerHisAspGluIle 20

Db 2388 GMAAAAGTGGTGTGATTAACCGGCACTGTACCCCTTGATTCAGAACCATGATGAATC 2447

QY 21 ThrIleGlnProGlnAspIleValMetValAspGluSerGlnThrGlyGluProGlyIle 40

Db 2448 ACATATCCAGCCAGAGACATAGTCATGTGTGATGAAGCCCAACTGGAGAACCCGGCTGG 2507

QY 41 LeuGIGlyGluLeuLysGlyLysThrGlyTYRPhenProAlaAsnTYRAlaGluLysIle 60

Db 2508 CTGGAGAGAGATTAAAGAAAGACAGGGGTGTTCTTCCTCAACTGTGACAGAAATATC 2567

QY 61 Pro 61

Db 2568 CCA 2570

RESULT 5
AA163825
ID AA163825 standard; cDNA; 3466 BP.

XX AA163825;

XX 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 33.

DE Human; antiarthritic; antirheumatic; antiproliferative; vasotropic;
KW cerebroprotective; nootropic; neuroprotective; antibacterial; vituicide;
KW fungicide; ophthalmological; cytostatic; immunosuppressive; nootropic;
KW neuroprotective; antiallergic; hepatotropic; antidiabetic;
KW antiinflammatory; antitumor; anticonvulsant; antibacterial;
KW antiparasitic; cardiant; gene therapy; cancer; immune disorder;
KW cardiovascular disorder; neurological disease; infection; human; ss.

XX Homo sapiens.

XX WO20015308-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001309.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-488781/53.
 PS P-PSDB; AAM43519.
 XX
 PT New isolated nucleic acids and polypeptides, useful for diagnosing,
 XX treating and/or preventing human diseases and disorders.
 XX
 PS Claim 1; SEQ ID NO 33; 664pp + Sequence Listing; English.
 XX
 CC The invention relates to human polynucleotides (AA163803-AA164012) and
 CC the encoded proteins (AAM43497-AAM43660) useful for preventing, treating
 CC or ameliorating medical conditions e.g. by protein or gene therapy. The
 CC genes were isolated from a range of human tissues disclosed in the
 CC specification. The nucleic acids, proteins, antibodies and (anti)agonists
 CC are useful in the diagnosis, treatment and prevention of: (a) cancer,
 CC e.g. breast and ovarian cancer and other cancers of the adrenal gland,

CC bone, bone marrow, breast, gastrointestinal tract, liver, lung, or
 CC urogenital; (b) immune disorders e.g. Addison's disease, allergies,
 CC autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
 CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
 CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections. Note: The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 3466 BP; 1056 A; 829 C; 861 G; 710 T; 0 U; 10 Other;
 SQ
 Alignment Scores:
 Pred. No.: 2,15e-34 Length: 3466
 Score: 333.00 Matches: 61
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_740_800 (1-61) x AA163825 (1-3466)
 QY 1 VallysValValTyrTrrrAgaAlaLeuTyrProphGgiuSerArGserHisAspGluile 20
 Db 2506 GTAAAGGTGGTATTACCGGCACTGTACCCCTTGTGATCCAGAGCCATGATGAATTC 2565
 QY 21 ThrileGlnProGlyAspIleValMetValAspGluSerGlnThrGlyuProGlyTrrp 40
 Db 2566 ACTATCCAGCCAGAGACATGTCATGTCGATGAATGAAGCCAACTGGAGAACCCGCTGG 2625
 QY 41 LeuGlyGlyGluLeuGlyGlyTrrpPheProAlaAsnTyrAlaGluIle 60
 Db 2626 CTTGAGAGAGATTTAAAGGAAGACAGGGTGTTCCTGCAACTATGACAGAGAAATTC 2685
 QY 61 Pro 61
 Db 2686 CCA 2688
 RESULT 6
 AAZ39009
 ID AAZ39009 standard; cDNA; 3723 BP.
 XX
 AC AAZ39009;
 XX
 DT 28-FEB-2000 (first entry)
 XX
 DE Mouse Esi coding sequence.
 XX
 KW Mouse; murine; Esi; Esi2; endocytosis; vesicular trafficking;
 KW regulation; actin cytoskeleton; detection; cancer; infection;
 KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
 KW antiproliferative; antiviral; ss.
 XX
 OS Mus sp.
 XX
 OS WO9955728-A2.
 PN
 PN 04-NOV-1999.
 PD
 PD 27-APR-1999; 99WO-CN000375.
 PF
 PF 27-APR-1998; 98CA-02230201.
 PR
 PR 05-FEB-1999; 99US-0118739P.
 PR
 PR (HSCR-) HSC RES & DEV LP.
 PA
 PA Egan SE, Wang W, Sengar A;
 XX
 XX WPI; 2000-052802/04.
 XX P-PSDB; AAY57444.
 DR
 DR New nucleic acid encoding Esi1 and 2 proteins, involved in regulation of

PT endocytosis, used e.g. for treating cancer or preventing viral infection.
XX
PS Claim 6; Page 40-42; 99pp; English.
XX
CC The present sequence encodes mouse Ese1. The present invention
CC specifically describes mammalian Ese1 and 2 proteins (I) and their splice
CC variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (I)
CC are involved in regulation of clathrin-mediated endocytosis (as a complex
CC with Esp1 protein), vesicular trafficking and actin cytoskeleton.
CC Generally (I) or its (ant)agonists, mimetics, fragments and inactive
CC mutants; (I)-specific antibodies (Ab); sequences antisense to the (I)
CC polynucleotide; agents that downregulate expression of Ese genes or
CC antagonists of an Ese binding partner are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of Ese1 is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (I) or Ab are used to suppress abnormal proliferation of cells that can
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-Esp1 complex, then binding dynam to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission
XX
SQ Sequence 3723 BP; 1065 A; 941 C; 1025 G; 692 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,35e-34 Length: 3723
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AA239009 (1-3723)
QY 1 ValLysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIle 20
DB 2212 GTAAAGTGTATATTACCGAGCGCTGTACCCCTTGAATCCGAAAGTCAGATGAGATC 2271
QY 21 ThrILeGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyrP 40
DB 2272 ACCATCCAGCCAGGAGATATATGTCATGTCGATGTAAGCAAGCCAGCTGAGAGCCAGGATGG 2331
QY 41 LeuGlyGlyGluLeuLysGlyTyrThrGlyTyrPheProAlaAsnTyrAlaGluLysIle 60
DB 2332 CTTGGAGGAGAGCTGAAGGGAAGACGGGATGTTCCCTGCAAACTATGAGAAAGATT 2391
QY 61 Pro 61
DB 2392 CCA 2394

RESULT 7
AA239008
ID AA239008 standard; cDNA; 5082 BP.
XX
AC AA239008;
XX
DT 28-FEB-2000 (first entry)
XX
DE Mouse Ese1 full length cDNA sequence.
XX
KW Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;
KW regulation; actin cytoskeleton; detection; cancer; infection;
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
KW antiproliferative; antiviral; ss.
XX
OS Mus sp.
XX
PN WO955728-A2.
XX

PD 04-NOV-1999.
XX
PF 27-APR-1999; 99WO-CA000375.
XX
PR 27-APR-1998; 98CA-02230201.
PR 05-FEB-1999; 99US-0118739P.
XX
PA (HSCR-) HSC RES & DEV LP.
XX
PI Egan SE, Wang W, Sengar A;
XX
DR WPI: 2000-052802/04.
DR P-PSDB; AA57444.
XX
PT New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.
XX
PS Claim 6; Page 38-40; 99pp; English.
XX
CC The present sequence encodes mouse Ese1. The present invention
CC specifically describes mammalian Ese1 and 2 proteins (I) and their splice
CC variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (I)
CC are involved in regulation of clathrin-mediated endocytosis (as a complex
CC with Esp1 protein), vesicular trafficking and actin cytoskeleton.
CC Generally (I) or its (ant)agonists, mimetics, fragments and inactive
CC mutants; (I)-specific antibodies (Ab); sequences antisense to the (I)
CC polynucleotide; agents that downregulate expression of Ese genes or
CC antagonists of an Ese binding partner are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of Ese1 is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (I) or Ab are used to suppress abnormal proliferation of cells that can
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-Esp1 complex, then binding dynam to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission
XX
SQ Sequence 5082 BP; 1363 A; 1253 C; 1391 G; 1075 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,46e-34 Length: 5082
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AA239008 (1-5082)
QY 1 ValLysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIle 20
DB 2470 GTAAAGTGTATATTACCGAGCGCTGTACCCCTTGAATCCGAAAGTCAGATGAGATC 2529
QY 21 ThrILeGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyrP 40
DB 2530 ACCATCCAGCCAGGAGATATATGTCATGTCGATGTAAGCAAGCCAGCTGAGAGCCAGGATGG 2589
QY 41 LeuGlyGlyGluLeuLysGlyTyrThrGlyTyrPheProAlaAsnTyrAlaGluLysIle 60
DB 2590 CTTGGAGGAGAGCTGAAGGGAAGACGGGATGTTCCCTGCAAACTATGAGAAAGATT 2649
QY 61 Pro 61
DB 2650 CCA 2652

RESULT 8
AA239025
ID AA239025 standard; cDNA; 5144 BP.
XX

AC AA239025;
 XX 28-FEB-2000 (first entry)
 XX Mouse EseqL coding sequence.
 DE Mouse EseqL coding sequence.
 KW Mouse; murine; EseqL; Eseq2; endocytosis; vesicular trafficking;
 KW regulation; actin cytoskeleton; detection; cancer; infection;
 KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
 KW antiproliferative; antiviral; ss.
 XX
 OS Mus sp.
 PN WO9955728-A2.
 XX 04-NOV-1999.
 PD
 XX 27-APR-1999; 99WO-CA000375.
 PF
 XX 27-APR-1998; 98CA-02230201.
 PR 05-FEB-1999; 99US-0118739P.
 XX
 PA (HSCR-) HSC RES & DEV LP.
 XX
 PI Egan SE, Wang W, Sengar A;
 XX
 XX WPI; 2000-052802/04.
 DR P-PSDB; AAY57449.
 XX
 PT New nucleic acid encoding Eseq1 and 2 proteins, involved in regulation of
 PT endocytosis, used e.g. for treating cancer or preventing viral infection.
 XX
 PS Claim 6; Page 59-62; 99pp; English.
 XX
 XX The present invention specifically describes mammalian Eseq1 and 2
 CC proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
 CC regulator of endocytosis). (I) are involved in regulation of clathrin-
 CC mediated endocytosis (as a complex with Eseq1 protein); vesicular
 CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
 CC mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
 CC sequences antisense to the (I) polynucleotide; agents that downregulate
 CC expression of Ese genes or antagonists of an Ese binding partner are used
 CC to treat diseases associated with undesirable endocytosis and resulting
 CC changes in cellular function. Particularly overexpression of Eseq1 is used
 CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
 CC administration of (I) is used to promote endocytosis of selected cells.
 CC (Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
 CC cells that can be stimulated to proliferate by a growth factor receptor;
 CC and similar compounds (also inactive Ese mutants) can be used to prevent
 CC viral infection. Endocytosis may also be regulated, in vivo or in cell
 CC cultures, by forming an Ese-Eseq1 complex, then binding dynamin to the
 CC complex. Generally conditions that can be treated include cancer;
 CC abnormal cell division or migration; viral infection; or abnormal
 CC receptor signalling, tissue development or synaptic transmission. The
 CC present sequence represents mouse EseqL coding sequence
 XX
 SQ Sequence 5144 BP; 1447 A; 1298 C; 1415 G; 984 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 3.51e-34 Length: 5144
 Score: 333.00 Matches: 61
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0
 US-09-720-934-2_COPY_740_800 (1-61) X AA239025 (1-5144)
 QY 1 VALYGVVALVATYTYRGAALALEUTYRPPHGLUSERARGSERHISAPGJULIE 20
 DB 2212 GTAAAGGTGTATATACCGAGCGCGTGTACCCCTTGATCCAGAGTACGATGATC 2271
 QY 21 ThrileGlnPrGJlYAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr 40

DB 2272 ACCATCCAGCCAGAGTATATAGCATGTGATGAAAGCCAGACTGAGGCCAGGATGG 2331
 QY 41 LengIyGlyGluLeuLysGlyTyrGnglyTyrPhProAlaSnYrAlaGlyLeIle 60
 DB 2332 CTGGAGAGAGAGCTGAAAGGAGAGCGGATGGTTCCTCCAACTATGAGAAAGATTT 2331
 QY 61 Pro 61
 DB 2392 CCA 2394
 RESULT 9
 ID AA234572 standard; cDNA; 5195 BP.
 XX
 AC AA234572;
 XX 01-FEB-2000 (first entry)
 DT
 XX
 DE Human SH3D1A cDNA clone 11.
 XX
 KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 239..386
 FT /*tag= a
 XX
 PN WO9953062-A2.
 XX
 XX 21-OCT-1999.
 PD
 XX 16-APR-1999; 99WO-US008371.
 PF
 XX 16-APR-1998; 98US-0082007P.
 PR
 XX (CEDA-) CEDARS STINAI HEALTH SYSTEM.
 PA
 PI Korenberg JR, Chen X;
 XX
 XX WPI; 1999-633829/54.
 DR P-PSDB; AAY32156.
 XX
 PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 XX
 PS Claim 2; Fig 10; 99pp; English.
 XX
 XX This is the nucleotide sequence of full-length cDNA (clone 11)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukaemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC platelet disorder on chromosome 21, low platelets in deletion for 21,
 CC association of gains in chromosome 21 with leukaemias, neural
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and
 CC treatment of a subject (including a prenatal subject) having

CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukaemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 XX
 SQ Sequence 5195 BP; 1592 A; 1148 C; 1237 G; 1218 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3,566-34 Length: 5195
 Score: 333.00 Matches: 61
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AA234570 (1-5195)

QY 1 ValysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIle 20
 DB 2456 GTAAAGTGGTGTATACCGGGCACTGTACCCCTTGATCCAGAGCCATGATGAATC 2515

QY 21 ThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr 40
 DB 2516 ACTATCCAGCCAGAGACATATGATGATGATGAAGCCAAACTGGAGAACCCGGCTGG 2575

QY 41 LeuGlyGlyGluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIle 60
 DB 2576 CTGGAGAGAAATTAAAGAAAGACAGGGGTGTTCCCTCCAAACTATGCAGAGAAATC 2635

QY 61 Pro 61
 DB 2636 CCA 2638

RESULT 10
 AA234570
 ID AA234570 standard; cDNA; 5199 BP.

AC AA234570;
 DT 01-FEB-2000 (first entry)
 XX
 DE Human SH3D1A cDNA.

XX SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH 208..3642
 FT /*tag= a
 XX

PN WO9553062-A2.

XX 21-OCT-1999.

PF 16-APR-1999; 99WO-US008371.

PR 16-APR-1998; 98US-0082007P.

XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.

XX Korenberg JR, Chen X;

XX WPI, 1999-633829/54.

XX P-PSDB; AAY32154.

XX Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 XX
 PS Claim 2; Fig 5; 99pp; English.

XX This is the nucleotide sequence of full-length cDNA corresponding to a
 CC novel human SH3 gene, termed the SH3D1A gene, that contributes to the
 CC development of platelets and the pathogenesis of leukaemia, both in
 CC general and in particular those involving the megakaryocytic lineage. The
 CC SH3D1A gene maps to the small candidate region for low platelets on
 CC chromosome 21. Sequencing of 5 different sizes of cDNA clone (see
 CC AA234570-74) suggests that at least 3 isoforms exist. The invention
 CC provides methods for the diagnosis and treatment of megakaryocytic
 CC abnormality, myeloproliferative disorder, platelet disorder, acute
 CC leukaemia, neural disorders, thrombocytopenia, platelet disorder on
 CC chromosome 21, low platelets in deletion for 21, association of gains in
 CC chromosome 21 with leukaemias, neural abnormalities, dysfunctions and
 CC disorders including brain malformations and corresponding cognitive
 CC dysfunctions, microcephaly, lissencephaly, and colpocephaly. Methods are
 CC also provided for: suppressing cells unable to regulate themselves;
 CC screening for a somatic alteration in the SH3D1A gene; monitoring the
 CC progress and adequacy of a treatment; monitoring tumour risk progress or
 CC disorder, platelet disorder or leukaemia; and treatment of a subject
 CC (including a prenatal subject) having megakaryocytic abnormality,
 CC myeloproliferative disorder, platelet disorder, leukaemia or neural
 CC disorder using a nucleic acid that expresses SH3D1A or its antisense
 CC nucleic acid
 XX
 SQ Sequence 5199 BP; 1615 A; 1152 C; 1214 G; 1218 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3,566-34 Length: 5199
 Score: 333.00 Matches: 61
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-09-720-934-2_COPY_740_800 (1-61) x AA234570 (1-5199)

QY 1 ValysValValTyrTyrArgAlaLeuTyrProPheGluSerArgSerHisAspGluIle 20
 DB 2425 GTAAAGTGGTGTATACCGGGCACTGTACCCCTTGATCCAGAGCCATGATGAATC 2484

QY 21 ThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr 40
 DB 2485 ACTATCCAGCCAGAGACATATGATGATGATGAAGCCAAACTGGAGAACCCGGCTGG 2544

QY 41 LeuGlyGlyGluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIle 60
 DB 2545 CTGGAGAGAAATTAAAGAAAGACAGGGGTGTTCCCTCCAAACTATGCAGAGAAATC 2604

QY 61 Pro 61
 DB 2605 CCA 2607

RESULT 11
 AA239024
 ID AA239024 standard; cDNA; 5738 BP.

AC AA239024;

DT 28-FEB-2000 (first entry)

XX Mouse Escl cDNA sequence.

XX Mouse; murine; Escl; Escl; endocytosis; vesicular trafficking;

XX regulation; actin cytoskeleton; detection; cancer; infection;

XX EH-domain and SH3-domain regulator of endocytosis; anticancer;

XX antiproliferative; antiviral; ss.
 OS Mus sp.
 XX
 XX WO955728-A2.
 XX
 XX 04-NOV-1999.

XX PF 27-APR-1999; 99WO-CA000375.
 XX PR 27-APR-1998; 98CA-02230201.
 XX PR 05-FEB-1999; 99US-0118739P.
 XX PA (HSCR-) HSC RES & DEV LP.
 XX PI Egan SE, Wang W, Sengar A;
 XX DR WPI: 2000-052802/04.
 XX DR P-PSDB; AAY57449.
 XX PT New nucleic acid encoding Esei and 2 proteins, involved in regulation of
 XX PT endocytosis, used e.g. for treating cancer or preventing viral infection.
 XX PS Claim 6; Page 56-59; 99pp; English.
 CC The present invention specifically describes mammalian Esei and 2
 CC proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
 CC regulator of endocytosis). (I) are involved in regulation of clathrin-
 CC mediated endocytosis (as a complex with Espis protein), vesicular
 CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
 CC mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
 CC expression of Ese genes or antagonists of an Ese binding partner are used
 CC to treat diseases associated with undesirable endocytosis and resulting
 CC changes in cellular function. Particularly overexpression of Esei is used
 CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
 CC administration of (I) is used to promote endocytosis of selected cells.
 CC (Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
 CC cells that can be stimulated to proliferate by a growth factor receptor;
 CC and similar compounds (also inactive Ese mutants) can be used to prevent
 CC viral infection. Endocytosis may also be regulated, in vivo or in cell
 CC cultures, by forming an Ese-Espis complex, then binding dynamin to the
 CC complex. Generally conditions that can be treated include cancer;
 CC abnormal cell division or migration; viral infection; or abnormal
 CC receptor signaling, tissue development or synaptic transmission. The
 CC present sequence represents mouse EseiL cDNA sequence
 CC
 CC SQ Sequence 5738 BP; 1556 A; 1456 C; 1613 G; 1113 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.02e-34 Length: 5738
 Score: 333.00 Matches: 61
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0
 US-09-720-934-2_COPY_740_800 (1-61) x AA239024 (1-5738)
 QY 1 VallysValValTyTyArgAlaLeuTyrrProphgUSeArgrSerHisAspGluile 20
 Db 2471 GTAAAGAGGTATATTATACGAGCGCTGTACCCCTTGATTCAGAAATCAGATGAGATC 2530
 QY 21 ThrllgInProGlyAspIleValMetValAspGluSeGlnThnGlyGluuproglyTTP 40
 Db 2531 ACCATCCAGCCAGGAGATATGTCATGTGGATGAAAGCCAGCTGGAGAGCCAGGATG 2590
 QY 41 LeuGlyGlyGluLeuLeuGlyHisThrGlyTTPPhProAlaAsnTyZAlaGluysile 60
 Db 2591 CTTGAGGAGAGCTGAAAGGAGAGACGGATGGTTCCTCCGAAACATATCAGAAAAAGATT 2650
 QY 61 Pro 61
 Db 2651 CCA 2653
 RESULT 12
 AA234571
 ID AA234571 standard; cDNA; 5458 BP.
 XX AA234571;
 AC

XX DT 01-FEB-2000 (first entry)
 XX DE Human SH3D1A cDNA clone 21.
 XX KW SH3D1A gene; human; Down's syndrome; leukemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW hematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT CDS 267..3929
 FT FT /**tag= a
 FT FT
 XX WO9953062-A2.
 XX PD 21-OCT-1999.
 XX PF 16-APR-1999; 99WO-US008371.
 XX PR 16-APR-1998; 98US-0082007P.
 XX PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 XX PI Korenberg JR, Chen X;
 XX DR WPI: 1999-633829/54.
 XX DR P-PSDB; AAY32155.
 XX PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 XX PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 XX PS Claim 2; Fig 6; 99pp; English.
 CC This is the nucleotide sequence of full-length cDNA (clone 21)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukaemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21 (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC association of gains in chromosome 21 with leukaemias, neural
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality; myeloproliferative
 CC disorder; haematopoietic disorder; platelet disorder or leukemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukaemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 CC
 CC SQ Sequence 5458 BP; 1671 A; 1201 C; 1307 G; 1279 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.81e-32 Length: 5458
 Score: 320.50 Matches: 61
 Percent Similarity: 92.428 Conservative: 0
 Best Local Similarity: 92.428 Mismatches: 0
 Query Match: 96.25% Indels: 5
 DB: 2 Gaps: 1
 US-09-720-934-2_COPY_740_800 (1-61) x AA234571 (1-5458)

QY 1 ValIysValIYrTYrArgAlaLeuTYrProPheGluSerArgSerHisAspGluIle 20
 Db 2484 GTAAAGGTGTGTATTACCGGCACTGTACCCCTTTGATCCAGAAAGCAGATGAATC 2543
 QY 21 ThrIleGlnProGlyAspIleValMet-----ValAspGluSerGlnThr 35
 Db 2544 ACTATCCAGCCAGAGACATAGTCATGTTAAAGGGGAAATGGTGGATGAAGCCAAACT 2603
 QY 36 GlyGluProGlyTYrPLeuGlyGlyGluLeuLYsGlyLysThrGlyTYrPheProAlaAsn 55
 Db 2604 GGAAACCCGCTGCTTGGAGGAAATTAAAGAAAGACAGGCTGTTCCCTGCAGAAAC 2663
 QY 56 TyrAlaGluLysIlePro 61
 Db 2664 TATGCAGAGAAATCCCA 2681

RESULT 13
 AAS84763
 ID AAS84763 standard; cDNA; 7435 BP.
 XX
 AC AAS84763;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #20567.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 XX
 PT P-PSDB; ABG20576.
 XX

Claim 1, SEQ ID NO 20567, 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant polynucleotide expression or biological activity. The polynucleotide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS84197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 7435 BP; 2190 A; 1690 C; 1829 G; 1726 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2,66e-32 Length: 7435
 Score: 320.50 Matches: 61
 Percent Similarity: 92.42% Conservative: 0
 Best Local Similarity: 92.42% Mismatches: 0
 Query Match: 96.25% Indels: 5
 DB: Gaps: 1

US-09-720-934-2_COPY_740_800 (1-61) x AAS84763 (1-7435)

QY 1 ValIysValIYrTYrArgAlaLeuTYrProPheGluSerArgSerHisAspGluIle 20
 Db 2499 GTAAAGGTGTGTATTACCGGCACTGTACCCCTTTGATCCAGAAAGCAGATGAATC 2558
 QY 21 ThrIleGlnProGlyAspIleValMet-----ValAspGluSerGlnThr 35
 Db 2559 ACTATCCAGCCAGAGACATAGTCATGTTAAAGGGGAAATGGTGGATGAAGCCAAACT 2618
 QY 36 GlyGluProGlyTYrPLeuGlyGlyGluLeuLYsGlyLysThrGlyTYrPheProAlaAsn 55
 Db 2619 GGAAACCCGCTGCTTGGAGGAAATTAAAGAAAGACAGGCTGTTCCCTGCAGAAAC 2678
 QY 56 TyrAlaGluLysIlePro 61
 Db 2679 TATGCAGAGAAATCCCA 2696

RESULT 14
 AAT39799
 ID AAT39799 standard; DNA; 2873 BP.
 XX
 AC AAT39799;
 XX
 DT 19-FEB-1998 (first entry)
 XX
 DE Human clone 65 gene.
 XX
 KW Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;
 KW cellular signalling element; cellular structural element; malignancy;
 KW protein identification; functional domain; protein screening;
 KW cellular signal transduction process; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 3..1532
 FT /*tag= a
 XX
 PN WO9631625-A1.
 PD 10-OCT-1996.
 XX
 PF 04-APR-1996; 96WO-US004454.
 PR 07-APR-1995; 95US-00417872.
 PR 03-APR-1996; 96US-00630915.
 XX
 PA (CYTO-) CYTOGEN CORP.
 PA (UYN-) UNIV NORTH CAROLINA.
 PI Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;
 XX
 DR WPI: 1996-465045/46.
 XX
 PT P-PSDB; AAM05399.
 XX

Identifying polypeptide(s) having specific functional domain (esp. SH3 domain) - comprises detecting selective binding to recognition unit, regardless of sequence homology.

PS Claim 76; Fig 58; 174pp; English.

CC AAT39786-T39803 represent novel human and mouse genes encoding Src-homology region 3 (SH3) domain containing proteins that can be used in the method of the invention. SH3 domain containing proteins play a role in signalling and structural elements of cells. The method of the invention is for identifying polypeptides containing functional domains of interest (especially SH3 domains). The method comprises contacting a multivalent recognition unit (RU) complex with a number of peptides and identifying polypeptides having a selective binding affinity for the RU complex. The method is based on functional similarities and does not rely on sequence similarities. Prior methods only gave limited success for identifying proteins which contain an SH3 domain due to the minimal sequence homology among known SH3 proteins. It has been found that small peptide RUs in multivalent form have reduced specificity for a given functional domain compared to monomer RUs. Multivalent RU complexes are particularly suited to screening for polypeptides containing functional domains that are similar to, but not identical in sequence to, the original target functional domain. The new method enables proteins having a common function to be identified. Identification of novel SH3 proteins will be useful for a better understanding of cell growth, malignancy, signal transduction processes, etc. New candidate drugs can be identified, and their specificities (e.g. pharmacological activities) can be assessed using the method of the invention.

CC XX Sequence 2873 BP; 903 A; 506 C; 623 G; 841 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	9.09e-19	Length:	2873
Score:	216.00	Matches:	36
Percent Similarity:	74.58%	Conservative:	8
Best Local Similarity:	61.02%	Mismatches:	15
Query Match:	64.86%	Indels:	0
DB:	2	Gaps:	0

US-09-720-934-2_COPY_740_800 (1-61) x AAT39799 (1-2873)

QY 3 ValValIYTYTYRAGAlaLeuTYRProBheGluSerArgSerHisAspGluIleThrIle 22

Db 228 TTGGTGAATTATAGACATTAACCCCTTGAAGCAAGAACCATGATGAGTGGTTT 287

QY 23 GlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTYRpleugly 42

Db 288 AATTCGTGAGATTAATTCAGGTTCATGAAAAAACCGAGAGAACCTGGTGGCTTAT 347

QY 43 GlyGluLeuIysGlyIysThrGlyTYRProBheProAlaAsnTYRAlaGluIysIlePro 61

Db 348 GGTAGTTTTCAGGAAATTTGGCTGTTTCATGCAATATATGTAGAAAAATGCCA 404

RESULT 15

AAZ39011

ID AAZ39011 standard; cDNA; 3593 BP.

XX AC AAZ39011;

DT 28-FEB-2000 (first entry)

XX Mouse E2e2 coding sequence.

DE Mouse

XX Mouse; murine; E2e1; E2e2; endocytosis; vesicular trafficking; regulation; actin cytoskeleton; detection; cancer; infection; E2e1 domain and SH3 domain regulator of endocytosis; anticancer; antiproliferative; antiviral; 88.

XX OS Mus sp.

XX PN WO955728-A2.

PD 04-NOV-1999.

XX 27-APR-1999; 99WO-CA000375.

XX 27-APR-1998; 96CA-02230201.

PR 05-FEB-1999; 99US-0118739P.

XX (HSCR-) HSC RES & DEV LP.

XX Bgan SE, Wang W, Sengar A;

XX WPI; 2000-052802/04.

DR P-PSDB; AAY57445.

PT New nucleic acid encoding E2e1 and 2 proteins, involved in regulation of endocytosis, used e.g. for treating cancer or preventing viral infection.

XX Claim 25; Page 46-48; 99pp; English.

CC The present sequence encodes mouse E2e2. The present invention specifically describes mammalian E2e1 and 2 proteins (I) and their splice variants (E2e = E2e-domain and SH3 domain regulator of endocytosis). (I) are involved in regulation of clathrin-mediated endocytosis (as a complex with E2e1 protein), vesicular trafficking and actin cytoskeleton. CC Generally (I) (or its (ant)agonists, mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab); sequences antisense to the (I) polynucleotide; agents that downregulate expression of E2e genes or antagonists of an E2e binding partner are used to treat diseases CC associated with undesirable endocytosis and resulting changes in cellular function. Particularly overexpression of E2e1 is used to block clathrin-mediated endocytosis in vivo or in cell cultures, while administration of (I) is used to promote endocytosis of selected cells. (ant)agonists of (I) or Ab are used to suppress abnormal proliferation of cells that can be stimulated to proliferate by a growth factor receptor; and similar CC compounds (also inactive E2e mutants) can be used to prevent viral infection. Endocytosis may also be regulated, in vivo or in cell CC cultures, by forming an E2e-E2e1 complex, then binding dynamin to the CC complex. Generally conditions that can be treated include cancer; abnormal cell division or migration; viral infection; or abnormal CC receptor signalling, tissue development or synaptic transmission

CC XX Sequence 3593 BP; 1171 A; 783 C; 850 G; 789 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.2e-18	Length:	3593
Score:	216.00	Matches:	36
Percent Similarity:	75.86%	Conservative:	8
Best Local Similarity:	62.07%	Mismatches:	14
Query Match:	64.86%	Indels:	0
DB:	3	Gaps:	0

US-09-720-934-2_COPY_740_800 (1-61) x AAZ39011 (1-3593)

QY 3 ValValIYTYTYRAGAlaLeuTYRProBheGluSerArgSerHisAspGluIleThrIle 22

Db 2155 TTGGTGAATTATAGACATTAACCCCTTGAAGCAAGAACCATGATGAGTGGTTT 2214

QY 23 GlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTYRpleugly 42

Db 2215 AGTTCGTGAGATTAATTCAGGTTCATGAAAAAACCTGAGAGAACCTGGTGGCTTAT 2274

QY 43 GlyGluLeuIysGlyIysThrGlyTYRProBheProAlaAsnTYRAlaGluIysIle 60

Db 2275 GGTAGTTTTCAGGAAATTTGGCTGTTTCATGCAATATATGTAGAAAAATGTG 2328

Search completed: July 1, 2004, 13:31:47

Job time : 199.286 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 11:41:07; Search time 182.112 Seconds
(without alignment)
1376.313 Million cell updates/sec

Title: US-09-720-934-2_COPY_1080_1138
Sequence: 1 AACVIGIMDYTAQNDEL.....GEVNGGVLPSPNVKLTLD 59

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delcp 6.0, Delcxt 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DRV=xlp
-O=/cgn2_1/USFTO.spool.B/US09720934/runat.30062004.064539.13422/app.query.fasta_1.1386
-DB=N.geneseq.290and4 -QFMT=fastap -SUFFIX=mg -MINMATCH=0.1 -LOOPLC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORMext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720934 @CGN 1.1 1596 -ext -runat.30062004.064539.13422 -NCPU=6 -ICPU=3
-NO MMAP -LARGEDUTRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: N Geneseq.29Jan04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002s:*\n7: geneseqn2003as:*\n8: geneseqn2003bs:*\n9: geneseqn2003cs:*\n10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	323	100.0	1389	2	AAT39795 Human SH3
2	323	100.0	2067	6	ABQ55007 Human ova
3	323	100.0	3231	2	AAZ34574 Human SH3
4	323	100.0	5195	2	AAZ34572 Human SH3
5	323	100.0	5199	2	AAZ34570 Human SH3
6	323	100.0	5458	2	AAZ34571 Human SH3
7	319	98.8	2874	5	AAZ34572 DNA encod
8	313	96.9	3723	3	AAZ39009 Mouse Bse

9	313	96.9	5092	3	AAZ39008 Mouse Ese
10	313	96.9	5144	3	AAZ39025 Mouse Ese
11	313	96.9	5738	3	AAZ39024 Mouse Ese
12	310	96.0	7435	5	AAZ39024 Mouse Ese
13	266	82.4	747	2	AAZ39796 Human SH3
14	266	82.4	2873	2	AAZ39799 Human SH3
15	266	82.4	3746	4	AAZ39799 Human SH3
16	266	82.4	4210	4	AAZ39799 Human SH3
17	266	82.4	4210	9	AAZ39799 Human SH3
18	266	82.4	4447	4	AAZ39799 Human SH3
19	266	82.4	5828	6	AAZ39799 Human SH3
20	266	82.4	6103	4	AAZ39799 Human SH3
21	266	81.4	4625	3	AAZ39010 Mouse Ese
22	263	81.4	4975	3	AAZ39027 Mouse Ese
23	263	81.4	6014	3	AAZ39026 Mouse Ese
24	242	74.9	3593	3	AAZ39011 Mouse Ese
25	191	59.1	3981	4	ABL01995 Human SH3
26	191	59.1	7225	4	ABL01994 Human SH3
27	160	49.5	393	3	AAZ39027 Mouse Ese
28	160	49.5	720	3	AAZ39027 Mouse Ese
29	160	49.5	1215	4	AAZ39027 Mouse Ese
30	160	49.5	1215	4	AAZ39027 Mouse Ese
31	160	49.5	1254	4	AAZ39027 Mouse Ese
32	160	49.5	1392	2	AAZ39027 Mouse Ese
33	160	49.5	1392	2	AAZ39027 Mouse Ese
34	160	49.5	1392	2	AAZ39027 Mouse Ese
35	160	49.5	1392	2	AAZ39027 Mouse Ese
36	160	49.5	1392	2	AAZ39027 Mouse Ese
37	160	49.5	1659	2	AAZ39027 Mouse Ese
38	160	49.5	1659	2	AAZ39027 Mouse Ese
39	160	49.5	1659	2	AAZ39027 Mouse Ese
40	160	49.5	1659	2	AAZ39027 Mouse Ese
41	160	49.5	2287	3	AAZ39027 Mouse Ese
42	160	49.5	2287	3	AAZ39027 Mouse Ese
43	160	49.5	2698	3	AAZ39027 Mouse Ese
44	160	49.5	2860	3	AAZ39027 Mouse Ese
45	160	49.5	2866	6	AAZ39027 Mouse Ese

ALIGNMENTS

RESULT 1	AAZ39795	standard; DNA; 1389 BP.
ID	AAZ39795	
AC	AAZ39795;	
XX		
DT	19-FEB-1998	(first entry)
XX		
DE	Human SH3P17 gene.	
XX		
KW	Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;	
KW	cellular signalling element; cellular structural element; malignancy;	
KW	protein identification; functional domain; protein screening;	
KW	cellular signal transduction process; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..1389
FT		/tag=a
FT		/product="SH3P17"
XX		
PN	W09631625-A1.	
XX		
PD	10-OCT-1996.	
XX		
PE	04-APR-1996;	96WO-US004454.
XX		
PR	07-APR-1995;	95US-00417872.
XX		
PR	03-APR-1996;	96US-00630915.
XX		
PA	(CYTO-) CYTOGEN CORP.	

(UYNC-) UNIV NORTH CAROLINA.
 PI Sparks AB, Hoffman N, Kay BK, Fowlkes DM, Mcconnell SJ;
 DR WPI; 1996-465045/46.
 DR P-PSDB; AAM05395.
 XX
 PT Identifying polypeptide(s) having specific functional domain (esp. SH3 domain) - comprises detecting selective binding to recognition unit, regardless of sequence homology;
 PT
 PS Claim 76; Fig 50; 174pp; English.
 XX
 CC AAT39786-T39803 represent novel human and mouse genes encoding Src-homology region 3 (SH3) domain containing proteins that can be used in the method of the invention. SH3 domain containing proteins play a role in signalling and structural elements of cells. The method of the invention is for identifying polypeptides containing functional domains of interest (especially SH3 domains). The method comprises contacting a multivalent recognition unit (RU) complex with a number of peptides and identifying polypeptides having a selective binding affinity for the RU complex. The method is based on functional similarities and does not rely on sequence similarities. Prior methods only gave limited success for identifying proteins which contain an SH3 domain due to the minimal sequence homology among known SH3 proteins. It has been found that small peptide RUs in multivalent form have reduced specificity for a given functional domain compared to monomer RUs. Multivalent RU complexes are particularly suited to screening for polypeptides containing functional domains that are similar to, but not identical in sequence to, the original target functional domain. The new method enables proteins having a common function to be identified. Identification of novel SH3 proteins will be useful for a better understanding of cell growth, malignancy, signal transduction processes, etc. New candidate drugs can be identified, and their specificities (e.g. pharmacological activities) can be assessed using the method of the invention
 CC
 XX SQ Sequence 1389 BP; 404 A; 370 C; 357 G; 258 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 2,01e-36 Length: 1389
 Score: 323.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_1080_1138 (1-59) x AAT39795 (1-1389)
 OY 1 AAlaValCyGgInValIleGlyNetyrAspYrTrhAlaGlnAsnAspAspGluLeuAla 20
 Db 1192 GGAGGTGTGCGAGGATGATGGATGATGAGTACGACGCGCCGAGATGACATGACTGGCC 1251
 OY 21 PheAsnLySGlyGlnIleIleAsnValIleuAsnLySGluAspProAspTrpTrpIySGly 40
 Db 1252 TTCACACAGGCGCCAGATCATCAACGTCCTCAACAGAGAGACCCCTGACTGTGGAAAGGA 1311
 OY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTrpValIyIleuThrTrhAsp 59
 Db 1312 GAAGTCATAGGACAGAGTGGGGCTTCCATCCATTAATGTGAAGTGCACAGAC 1368
 RESULT 2
 AB055007
 ID AB055007 standard; cDNA; 2067 BP.
 AC AB055007;
 XX
 XX 22-AUG-2002 (first entry)
 DE Human ovarian antigen HVBVX28 cDNA, SEQ ID NO:887.
 XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast; cancer; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;

KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; chromosome 21q22.1-22.2;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN W020020677-A1.
 XX
 PD 03-JAN-2002.
 XX
 PF 07-JUN-2001; 2001WO-US018569.
 XX
 PR 07-JUN-2000; 2000US-0209467P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI; 2002-147878/19.
 DR P-PSDB; ABP41930.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides, useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian cancer), immune disorders, cardiovascular disorders and neurological diseases.
 PT
 PS Claim 1; SEQ ID NO 887; 2922pp; English.
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-ABP43228) and to cDNAs encoding them (AB054131-AB056305), and also encompasses polypeptides 90% identical and polynucleotides 95% identical to the sequences of the invention. The invention additionally relates to recombinant vectors and host cells comprising human ovarian antigen CC polynucleotides, antibodies against human ovarian antigens, and the use CC of ovarian antigen polynucleotides and polypeptides in diagnosing, CC treating, prognosing or preventing various ovary and/or breast-related CC disorders. Such conditions include ovarian cancer and breast cancer, and CC metastatic tumours of ovarian or breast origin, reproductive system CC disorders (e.g., infertility, disorders of pregnancy, anovulation, CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and CC vaginitis), immune disorders (e.g., congenital and acquired CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), CC blood-related disorders (e.g., anaemia), cardiovascular disorders, CC respiratory disorders, neurological disorders, gastrointestinal disorders CC and urinary system disorders. Ovarian antigen polypeptides and CC polynucleotides may also be used in screening for compounds which CC modulate ovarian antigen expression or activity. The polynucleotides may CC further be used for gene therapy, chromosome mapping, in the CC identification of individuals and in forensic analysis, and the CC polypeptides may be used as food additives or to prepare antibodies CC useful in disease diagnosis, drug targeting and phenotyping. The present CC sequence represents cDNA encoding a human ovarian antigen of the CC invention. Note: The sequence data for this patent did not form part of CC the printed specification, but was obtained in electronic format directly CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2067 BP; 614 A; 426 C; 418 G; 605 T; 0 U; 4 Other;
 SQ
 Alignment Scores:
 Pred. No.: 3.32e-36 Length: 2067
 Score: 323.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

CC This is the nucleotide sequence of full-length cDNA (clone 11)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clones from foetal brain (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis,
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC platelet disorder on chromosome 21, low platelets in deletion for 21,
 CC association of gains in chromosome 21 with leukaemias, neural
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukaemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid

XX Sequence 5195 BP; 1592 A; 1148 C; 1237 G; 1218 T; 0 U; 0 Other;

SQ

Alignment Scores:

Pred. No.:	1.07e-35	Length:	5195
Score:	323.00	Matches:	59
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) x AA234572 (1-5195)

QY 1 AAlaValCyGGInValIIleGlyMeCTyrAspTYrThrAlaGlnAsnAspAspGluLeuAla 20

Db 3689 GCAGGTGCCAGGTAATGGAGTGTACGACTACACCCGCCAGATGACGATGAGCTGGCC 3748

QY 21 PheAsnLysGlyGlnIleleAsnValLeuAsnLysGluAspProAspTyrTrpLysGly 40

Db 3749 TTCAACAAGGGCCAGATCATCAACGTCTCCAAACAAGAGGAGACCTGACTGGTGAAGAAGA 3808

QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTYrVallyLeuThrThrAsp 59

Db 3809 GAAGTCATGACACAGTGGGGCTCTCCATCCAAATTATGTGAAGCTGACCAAGAC 3865

RESULT 5

AA234570

ID AA234570 standard; cDNA; 5199 BP.

AC AA234570;

DT 01-FEB-2000 (first entry)

XX Human SH3D1A cDNA.

DE SH3D1A gene; human; Down's syndrome; leukaemia; cancer;

XX megakaryocytic abnormality; myeloproliferative disorder;

KW platelet disorder; neural disorder; thrombocytopenia;

KM haematopoietic disorder; cognitive dysfunction; microcephaly;

XX lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.

OS Homo sapiens.

XX Homo sapiens.

FH Key Location/Qualifiers

FT CDS 208..3642

XX /tag= a

XX MO9953062-A2.

XX PD 21-OCT-1999.

XX

PF 16-APR-1999; 99MO-US008371.

XX

PR 16-APR-1998; 98US-0082007P.

XX

PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.

XX

PI Korenberg JR, Chen X;

XX

DR WPI; 1999-633829/54.

XX

PT P-PSDB; AAT32154.

XX

PS Nucleic acid from the human SH3D1A gene and its products, useful for the
 diagnosis and treatment of myeloproliferative disorders and leukemia.

XX

Claim 2; Fig 5; 99P; English.

XX

CC This is the nucleotide sequence of full-length cDNA corresponding to a
 CC novel human SH3 gene, termed the SH3D1A gene, that contributes to the
 CC development of platelets and the pathogenesis of leukaemias, both in
 CC general and in particular those involving the megakaryocytic lineage. The
 CC SH3D1A gene maps to the small candidate region for low platelets on
 CC chromosome 21. Sequencing of 5 different sizes of cDNA clones (see
 CC AA234570-74) suggests that at least 3 isoforms exist. The invention
 CC provides methods for the diagnosis and treatment of megakaryocytic
 CC abnormality, myeloproliferative disorder, platelet disorder, acute
 CC leukaemia, neural disorders, thrombocytopenia, platelet disorder on
 CC chromosome 21, low platelets in deletion for 21, association of gains in
 CC chromosome 21 with leukaemias, neural abnormalities, dysfunctions and
 CC disorders including brain malformations and corresponding cognitive
 CC dysfunctions, microcephaly, lissencephaly, and colpocephaly. Methods are
 CC also provided for: suppressing cells unable to regulate themselves;
 CC screening for a somatic alteration in the SH3D1A gene; monitoring the
 CC progress and adequacy of a treatment; monitoring tumour risk progress or
 CC megakaryocytic abnormality, myeloproliferative disorder, haematopoietic
 CC disorder, platelet disorder or leukaemia; and treatment of a subject
 CC (including a prenatal subject) having megakaryocytic abnormality,
 CC myeloproliferative disorder, platelet disorder, leukaemia or neural
 CC disorder using a nucleic acid that expresses SH3D1A or its antisense
 CC nucleic acid

XX Sequence 5199 BP; 1615 A; 1152 C; 1214 G; 1218 T; 0 U; 0 Other;

SQ

Alignment Scores:

Pred. No.:	1.07e-35	Length:	5199
Score:	323.00	Matches:	59
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) x AA234570 (1-5199)

QY 1 AAlaValCyGGInValIIleGlyMeCTyrAspTYrThrAlaGlnAsnAspAspGluLeuAla 20

Db 3445 GCAGGTGCCAGGTAATGGAGTGTACGACTACACCCGCCAGATGACGATGAGCTGGCC 3504

QY 21 PheAsnLysGlyGlnIleleAsnValLeuAsnLysGluAspProAspTyrTrpLysGly 40

Db 3505 TTCAACAAGGGCCAGATCATCAACGTCTCCAAACAAGAGGAGACCTGACTGGTGAAGAAGA 3564

QY 41 GluValAsnGlyGlnValGlyLeuPheProSerAsnTYrVallyLeuThrThrAsp 59

Db 3565 GAAGTCATGACACAGTGGGGCTCTCCATCCAAATTATGTGAAGCTGACCAAGAC 3621

RESULT 6

AA234571

ID AA234571 standard; cDNA; 5458 BP.

AC AA234571;

XX

DT 01-FEB-2000 (first entry)

XX

DE Human SH3D1A cDNA clone 21.
 XX
 KW SH3D1A gene; human; Down's syndrome; leukemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT CDS 267..3929
 FT /*tag= a
 XX
 PN W09953062-A2.
 XX
 PD 21-OCT-1999.
 XX
 PF 16-APR-1999; 99WC-US008371.
 XX
 PR 16-APR-1998; 98US-0082007P.
 XX
 PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 XX
 PI Korenberg JR, Chen X;
 XX
 DR WPI; 1999-633829/54.
 DR P-PSDB; AAY32155.
 XX
 PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 PS
 PS Claim 2; Fig 8; 99p; English.
 XX
 CC This is the nucleotide sequence of full-length cDNA (clone 21)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see A4234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC association of gains in chromosome 21, low platelets in deletion for 21,
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 XX
 SQ Sequence 5458 BP; 1671 A; 1201 C; 1307 G; 1279 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.14e-35 Length: 5458
 Score: 323.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0
 US-09-720-934-2_COPY_1080_1138 (1-59) x AA234571 (1-5458)
 QY 1 A1AVal1CysG1nVal1IleG1MetTyrAspTyrTrpAlaGlnAsnAspAspG1uLeuAla 20
 Db 3732 GCAAGTGTCCAGGATGGATGGATGACATACCGCGCGAAGTGAAGATGAGTGGCC 3791

QY 21 PheAsnIysG1nIleIleAsnValLeuAsnIysG1uAAspProAspTyrTrpIysG1y 40
 Db 3792 TTCAACAAGGCGCCAGATCATCAAGTCTCTACAGAAGAGACCTGATGTGGAAAGCA 3851
 QY 41 GluValAsnG1nVal1GlyLeuPheProSerAsnTyrVal1LysLeuThrThraSp 59
 Db 3852 GAAGTCATATGACAAAGTGGGCTCTTCCATCCATATATGTGAAGCTACCCACAGAC 3908
 RESULT 7
 AAS84762
 ID AAS84762 standard; cDNA; 2874 BP.
 XX
 AC AAS84762;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #20566.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR P-PSDB; ABG20575.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID NO 20566; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2874 BP; 940 A; 679 C; 703 G; 552 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.9e-35 Length: 2874

Score: 319.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 98.76%
DB: 5

Matches: 58
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x AAS84762 (1-2874)

QY 2 ValAysglValValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAlaPhe 21

DB 1855 GTGTGCCAGGTGATGGATGTACGACTACCGCGAGATGACGATGAGCTGGCTTC 1914

QY 22 AsnLysGlyGlnIleIleAsnValLeuAsnLysGlnAspProAspTyrPlyGlyGlu 41

DB 1915 AACAGGCGCCAGATCAACAGCTCTCAACAGAGGACCCCTGCTGTGGAAAGGAA 1974

QY 42 ValAsnGlyGlnValAlaGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59

DB 1975 GTCAATGGACAGTGGGCTCTTCCCATCCATTATGTGAAGTCGACCAAGAC 2028

RESULT 8
AAZ39009 standard; cDNA; 3723 BP.

XX AAZ39009;

XX 28-FEB-2000 (first entry)

XX Mouse Esel coding sequence.

XX Mouse Esel coding sequence.

XX Mouse; murine; Esel; Esee2; endocytosis; vesicular trafficking;

XX regulation; actin cytoskeleton; detection; cancer; infection;

XX EH-domain and SH3-domain regulator of endocytosis; anticancer;

XX antiproliferative; antiviral; ss.

XX Mus sp.

XX WO9955728-A2.

XX 04-NOV-1999.

XX 27-APR-1999; 99WO-CA000375.

XX 27-APR-1998; 98CA-02230201.

XX 05-FEB-1999; 99US-0118739P.

XX (HSCR-) HSC RES & DEV LP.

XX Egan SE, Wang W, Sengar A;

XX WPI; 2000-052802/04.

XX P-PSDB; AAY57444.

XX New nucleic acid encoding Esel and 2 proteins, involved in regulation of

XX endocytosis, used e.g. for treating cancer or preventing viral infection.

XX Claim 6; Page 40-42; 99pp; English.

XX The present sequence encodes mouse Esel. The present invention

XX specifically describes mammalian Esel and 2 proteins (1) and their splice

XX variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (1)

XX are involved in regulation of clathrin-mediated endocytosis (as a complex

XX with Esp15 protein), vesicular trafficking and actin cytoskeleton.

XX Generally (1) (or its (ant)agonists, mimetics, fragments and inactive

XX mutants); (1)-specific antibodies (Ab); sequences antisense to the (1)

XX polynucleotide; agents that downregulate expression of Ese genes or

XX antagonists of an Ese binding partner are used to treat diseases

XX associated with undesirable endocytosis and resulting changes in cellular

XX function. Particularly overexpression of Esel is used to block clathrin-

XX mediated endocytosis in vivo of in cell cultures, while administration of

XX (1) is used to promote endocytosis of selected cells. (Ant)agonists of

XX (1) or Ab are used to suppress abnormal proliferation of cells that can

XX be stimulated to proliferate by a growth factor receptor; and similar

CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-Esp15 complex, then binding dynamin to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission

QY Sequence 3723 BP; 1065 A; 941 C; 1025 G; 692 T; 0 U; 0 Other;

DB Alignment Scores: 1,92e-34 Length: 3723

QY Pred. No.: 313.00 Matches: 57

Score: 313.00 Conservative: 2

Percent Similarity: 100.00% Mismatches: 0

Best Local Similarity: 96.61% Indels: 0

Query Match: 96.90% Gaps: 0

DB: 3

US-09-720-934-2_COPY_1080_1138 (1-59) x AAZ39009 (1-3723)

QY 1 AlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20

DB 3445 GCAGTGTGCCAGGTGATCGGATGTACGATTACACCGCCGAGACGATGACGAACTAGCC 3504

QY 21 PheAsnLysGlyGlnIleIleAsnValLeuAsnLysGlnAspProAspTyrPlyGlyGlu 40

DB 3505 TTCAGCAAGGCCAGATCATCAACGCTCTCAACAGAGGAGACCGGACTGTGAAAGGA 3564

QY 41 GluValAsnGlyGlnValAlaGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59

DB 3565 GAAGTCAGTGGGACAGTGGGCTCTTCCCATCCATTATGTGAAGTCGACCAAGAC 3621

RESULT 9
AAZ39008 standard; cDNA; 5082 BP.

XX AAZ39008;

XX 28-FEB-2000 (first entry)

XX Mouse Esel full length cDNA sequence.

XX Mouse; murine; Esel; Esee2; endocytosis; vesicular trafficking;

XX regulation; actin cytoskeleton; detection; cancer; infection;

XX EH-domain and SH3-domain regulator of endocytosis; anticancer;

XX antiproliferative; antiviral; ss.

XX Mus sp.

XX WO9955728-A2.

XX 04-NOV-1999.

XX 27-APR-1999; 99WO-CA000375.

XX 27-APR-1998; 98CA-02230201.

XX 05-FEB-1999; 99US-0118739P.

XX (HSCR-) HSC RES & DEV LP.

XX Egan SE, Wang W, Sengar A;

XX WPI; 2000-052802/04.

XX P-PSDB; AAY57444.

XX New nucleic acid encoding Esel and 2 proteins, involved in regulation of

XX endocytosis, used e.g. for treating cancer or preventing viral infection.

XX Claim 6; Page 38-40; 99pp; English.

XX The present sequence encodes mouse Esel. The present invention

XX specifically describes mammalian Esel and 2 proteins (1) and their splice

XX variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (1)

XX are involved in regulation of clathrin-mediated endocytosis (as a complex

CC with BspI5 protein), vesicular trafficking and actin cytoskeleton.
CC Generally (I) (or its (ant)agonists, mimetics, fragments and inactive
CC mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
CC polynucleotide; agents that downregulate expression of Ese genes or
CC antagonists of an Ese binding partner, are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of Ese1 is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (I) is used to promote endocytosis of selected cells. (ant)agonists of
CC (I) or Ab are used to suppress abnormal proliferation of cells that can
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-BspI5 complex, then binding dynamin to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission
XX

SQ Sequence 5082 BP; 1363 A; 1253 C; 1391 G; 1075 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2, 84e-34 Length: 5082
Score: 313.00 Matches: 57
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 96.61% Mismatches: 0
Query Match: 96.90% Indels: 0
DB: 3 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x AA239008 (1-5082)

QY 1 AAlaValCysGlnValIleGlyMetYrAspTYrThrAlaGlnAsnAspGluLeuAla 20
DB 3702 GGAAGTGTCCAGGTGATCGGGATGTACGATTACACCGCCAGAACATATGCACTAGCC 3761
QY 21 PheAsnIysGlyGlnIleIleAsnValIleuAsnIysGlyAspProAspTrrpIysGly 40
DB 3762 TTCAGCAAGGCGCAGATCATCAACGCTCCACCAAGAGAGACCGGACTGGTGAAGA 3821
QY 41 GluValaAsnGlyGlnValaGlyLeuPheProSerAsnTYrValIysLeuThrThrasp 59
DB 3822 GAAGTCAGTGGCGCAAGTTGGCTCTTCCATCCCAATTATGTAAAGCTGACCAAGAC 3878

RESULT 10
AA239025
ID AA239025 standard; cDNA; 5144 BP.

XX AA239025;

DT 28-FEB-2000 (first entry)

XX Mouse Ese1L coding sequence.

XX Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;
KW regulation; actin cytoskeleton; detection; cancer; infection;
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
XX antiproliferative; antiviral; ss.

XX Mus sp.

XX MO9955728-A2.

PD 04-NOV-1999.

XX 27-APR-1999; 99WO-CA000375.

XX 27-APR-1998; 98CA-02230201.
PR 05-FEB-1999; 99US-0118739P.

XX (HSCR-) HSC RES & DEV LP.

XX Egan SE, Wang W, Sengar A;

XX WPI; 2000-052802/04.

DR P-PSDB: AA575449.

XX New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.
XX

PS Claim 6; Page 55-62; 99pp; English.

CC The present invention specifically describes mammalian Ese1 and 2
CC proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
CC regulator of endocytosis). (I) are involved in regulation of clathrin-
CC mediated endocytosis (as a complex with BspI5 protein), vesicular
CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
CC mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
CC sequences antisense to the (I) polynucleotide; agents that downregulate
CC expression of Ese genes or antagonists of an Ese binding partner are used
CC to treat diseases associated with undesirable endocytosis and resulting
CC changes in cellular function. Particularly overexpression of Ese1 is used
CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
CC administration of (I) or Ab are used to suppress abnormal proliferation of
CC cells that can be stimulated to proliferate by a growth factor receptor;
CC and similar compounds (also inactive Ese mutants) can be used to prevent
CC viral infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-BspI5 complex, then binding dynamin to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission. The
XX present sequence represents mouse Ese1L coding sequence

SQ Sequence 5144 BP; 1447 A; 1298 C; 1415 G; 984 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2, 89e-34 Length: 5144
Score: 313.00 Matches: 57
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 96.61% Mismatches: 0
Query Match: 96.90% Indels: 0
DB: 3 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x AA239025 (1-5144)

QY 1 AAlaValCysGlnValIleGlyMetYrAspTYrThrAlaGlnAsnAspGluLeuAla 20
DB 3445 GGAAGTGTCCAGGTGATCGGGATGTACGATTACACCGCCAGAACATATGCACTAGCC 3504

QY 21 PheAsnIysGlyGlnIleIleAsnValIleuAsnIysGlyAspProAspTrrpIysGly 40
DB 3505 TTCAGCAAGGCGCAGATCATCAACGCTCCACCAAGAGAGACCGGACTGGTGAAGA 3564

QY 41 GluValaAsnGlyGlnValaGlyLeuPheProSerAsnTYrValIysLeuThrThrasp 59
DB 3565 GAAGTCAGTGGCGCAAGTTGGGCTCTTCCATCCCAATTATGTAAAGCTGACCAAGAC 3621

RESULT 11
AA239024
ID AA239024 standard; cDNA; 5738 BP.

XX AA239024;

DT 28-FEB-2000 (first entry)

XX Mouse Ese1L cDNA sequence.

XX Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;
KW regulation; actin cytoskeleton; detection; cancer; infection;
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
XX antiproliferative; antiviral; ss.

XX Mus sp.

XX MO9955728-A2.

XX 04-NOV-1999.

XX 27-APR-1999; 99WO-CA000375.
XX 27-APR-1998; 98CA-02230201.
XX 05-FEB-1999; 99US-0118739P.
XX (HSCR-) HSC RES & DEV LP.
XX Egan SE, Wang W, Sengar A;
XX WPI; 2000-052802/04.
XX P-PSDB; AAY57449.
XX New nucleic acid encoding Ees1 and 2 proteins, involved in regulation of
XX endocytosis, used e.g. for treating cancer or preventing viral infection.
XX
XX Claim 6; Page 56-59; 99p; English.
XX
XX The present invention specifically describes mammalian Ees1 and 2
XX proteins (I) and their splice variants (Ees = EH-domain and SH3-domain
XX regulator of endocytosis). (I) are involved in regulation of clathrin-
XX mediated endocytosis (as a complex with Ees1 protein), vesicular
XX trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
XX mimetics, fragments and inactive mutants); (II)-specific antibodies (Ab);
XX sequences antisense to the (I) polynucleotide; agents that downregulate
XX expression of Ees genes or antagonists of an Ees binding partner are used
XX to treat diseases associated with undesirable endocytosis and resulting
XX changes in cellular function. Particularly overexpression of Ees1 is used
XX to block clathrin-mediated endocytosis in vivo or in cell cultures, while
XX administration of (I) is used to promote endocytosis of selected cells.
XX (Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
XX cells that can be stimulated to proliferate by a growth factor receptor;
XX and similar compounds (also inactive Ees mutants) can be used to prevent
XX viral infection. Endocytosis may also be regulated, in vivo or in cell
XX cultures, by forming an Ees-Ees1s complex, then binding dynamin to the
XX complex. Generally conditions that can be treated include cancer;
XX abnormal cell division or migration; viral infection; or abnormal
XX receptor signalling, tissue development or synaptic transmission. The
XX present sequence represents mouse Ees1L cDNA sequence
XX
XX SQ Sequence 5738 BP; 1556 A; 1456 C; 1613 G; 1113 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 3.32e-34 Length: 5738
XX Score: 313.00 Matches: 57
XX Percent Similarity: 100.00% Conservative: 2
XX Best Local Similarity: 96.61% Mismatches: 0
XX Query Match: 96.90% Indels: 0
XX DB: Gaps: 0
XX
XX US-09-720-934-2_COPY_1080_1138 (1-59) x AAZ39024 (1-5738)
XX
XX QY 1 AAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
XX DB 3704 GCAGGTGCGCAGGTGATCGGATGTAACGATACACCGCCCAAGATACCAACTAGCC 3763
XX QY 21 PheAsnLysGlyGlnIleLeuValLeuAsnLysGluAspProAspTyrTrpLysGly 40
XX DB 3764 TTCACCAAGGCGCAGATCATCAACGCTCTCAACAGAGAGACCCGACCTGGTGGAAAGCA 3823
XX QY 41 GluValaLsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59
XX DB 3824 GAAGTCAGTGGCGAAGTGGGTTCTTCCATCCAAATATGTAAGCTGACCAAGAC 3880
XX
XX RESULT 12
XX AAS84763
XX AAS84763 standard; cDNA; 7435 BP.
XX AAS84763;
XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #20567.

XX XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX
XX Dermanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX P-PSDB; ABG20576.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensic, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 1; SEQ ID NO 20567; 103p; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful for treating disorders
XX of sites expressing (II). (I) and (II) are useful for identifying disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensic, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pcc_sequences
XX
XX SQ Sequence 7435 BP; 2190 A; 1690 C; 1829 G; 1726 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1.24e-33 Length: 7435
XX Score: 310.00 Matches: 59
XX Percent Similarity: 98.33% Conservative: 0
XX Best Local Similarity: 98.33% Mismatches: 0
XX Query Match: 95.98% Indels: 1
XX DB: Gaps: 0
XX
XX US-09-720-934-2_COPY_1080_1138 (1-59) x AAS84763 (1-7435)
XX
XX QY 1 AAlaValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAsnAspAspGluLeuAla 20
XX DB 3758 GCAGGTGCGCAGGTGATCGGATGTAACGATACACCGCCCAAGATACCAACTAGCC 3817
XX QY 21 PheAsnLys-GlyGlnIleLeuValLeuAsnLysGluAspProAspTyrTrpLysGly 40
XX DB 3818 TTCACCAAGGCGCAGATCATCAACGCTCTCAACAGAGAGACCCGACCTGGTGGAAAGC 3877
XX QY 40 YGluValaLsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThrAsp 59

DB 3878 AGAAGTCATGACAGTGGGCTCTTCCATCCATTATGTGAAGTGCACGAGAC 3935

RESULT 13

AAT39796

ID AAT39796 standard; DNA; 747 BP.

XX

AC AAT39796;

XX

DT 19-FEB-1998 (first entry)

XX

DE Human SH3P18 gene.

XX

KM Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;

KW cellular signalling element; cellular structural element; malignancy;

KM protein identification; functional domain; protein screening;

KW cellular signal transduction process; ss.

XX

OS Homo sapiens.

XX

FH Key

FT 1. 747

FT CDS /tag= a

FT /product= "SH3P18"

XX

PN MO9631625-A1.

XX

PD 10-OCT-1996.

XX

PF 04-APR-1996; 96MO-US004454.

XX

PR 07-APR-1995; 95US-00417872.

PR 03-APR-1996; 96US-00630915.

XX

PA (CYTO-) CYTOGEN CORP.

PA (UNYC-) UNIV NORTH CAROLINA.

XX

PI Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;

XX

DR WPI; 1996-465045/46.

DR P-PSDB; AAW05396.

XX

PT Identifying polypeptide(s) having specific functional domain (esp. SH3

PT domain) - comprises detecting selective binding to recognition unit,

PT regardless of sequence homology.

XX

PS Claim 76; Fig 52; 174pp; English.

XX

CC AAT39786-139803 represent novel human and mouse genes encoding Src-

CC homology region 3 (SH3) domain containing proteins that can be used in

CC the method of the invention. SH3 domain containing proteins play a role

CC in signalling and structural elements of cells. The method of the

CC invention is for identifying polypeptides containing functional domains

CC of interest (especially SH3 domains). The method comprises contacting a

CC multivalent recognition unit (RU) complex with a number of peptides and

CC identifying polypeptides having a selective binding affinity for the RU

CC complex. The method is based on functional similarities and does not rely

CC on sequence similarities. Prior methods only gave limited success for

CC identifying proteins which contain an SH3 domain due to the minimal

CC sequence homology among known SH3 proteins. It has been found that small

CC peptide RUs in multivalent form have reduced specificity for a given

CC functional domain compared to monomer RUs. Multivalent RU complexes are

CC particularly suited to screening for polypeptides containing functional

CC domains that are similar to, but not identical in sequence to, the

CC original target functional domain. The new method enables proteins having

CC a common function to be identified. Identification of novel SH3 proteins

CC signal transduction processes, etc. New candidate drugs can be

CC identified, and their specificities (e.g. pharmacological activities) can

CC be assessed using the method of the invention

XX

SQ Sequence 747 BP; 244 A; 133 C; 186 G; 184 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.41e-28 Length: 747

Score: 266.00 Matches: 44

Percent Similarity: 91.38% Conservative: 9

Best Local Similarity: 75.86% Mismatches: 5

Query Match: 82.35% Indels: 0

DB: 2 Gaps: 0

US-09-720-934-2_COPY_1080_1138 (1-59) x AAT39796 (1-747)

QY 2 ValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnSerAspGluLeuAlaPhe 21

DB 553 GTATGTACAGGTGATTCGTATGATGACTATGCAGCAATATATGAAATGAGCTCAGTTTC 612

QY 22 AsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrTyrLysGlyGlu 41

DB 613 TCCAGGGACACACTCATTAATGTTATGACAAAGATATCTGATTGGTGGCAAGAGAG 672

QY 42 ValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrAsp 59

DB 673 ATCAAGGGGTGACGTGGTCTCTTCCCTCAAACTACGTTAAGATGACGACGAGAC 726

RESULT 14

AAT39799

ID AAT39799 standard; DNA; 2873 BP.

XX

AC AAT39799;

XX

DT 19-FEB-1998 (first entry)

XX

DE Human clone 65 gene.

XX

KM Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;

KW cellular signalling element; cellular structural element; malignancy;

KW protein identification; functional domain; protein screening;

KW cellular signal transduction process; ss.

XX

OS Homo sapiens.

XX

FH Key

FT 3. 1532

FT CDS /tag= a

FT

XX

PN MO9631625-A1.

XX

PD 10-OCT-1996.

XX

PF 04-APR-1996; 96MO-US004454.

XX

PR 07-APR-1995; 95US-00417872.

PR 03-APR-1996; 96US-00630915.

XX

PA (CYTO-) CYTOGEN CORP.

PA (UNYC-) UNIV NORTH CAROLINA.

XX

PI Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;

XX

DR WPI; 1996-465045/46.

DR P-PSDB; AAW05399.

XX

PT Identifying polypeptide(s) having specific functional domain (esp. SH3

PT domain) - comprises detecting selective binding to recognition unit,

PT regardless of sequence homology.

XX

PS Claim 76; Fig 58; 174pp; English.

XX

CC AAT39786-139803 represent novel human and mouse genes encoding Src-

CC homology region 3 (SH3) domain containing proteins that can be used in

CC the method of the invention. SH3 domain containing proteins play a role

CC in signalling and structural elements of cells. The method of the

CC invention is for identifying polypeptides containing functional domains

CC of interest (especially SH3 domains). The method comprises contacting a

CC multivalent recognition unit (RU) complex with a number of peptides and

CC identifying polypeptides having a selective binding affinity for the RU

CC complex. The method is based on functional similarities and does not rely
CC on sequence similarities. Prior methods only gave limited success for
CC identifying proteins which contain an SH3 domain due to the minimal
CC sequence homology among known SH3 proteins. It has been found that small
CC peptide RUs in multivalent form have reduced specificity for a given
CC functional domain compared to monomer RUs. Multivalent RU complexes are
CC particularly suited to screening for polypeptides containing functional
CC domains that are similar to, but not identical in sequence to, the
CC original target functional domain. The new method enables proteins having
CC a common function to be identified. Identification of novel SH3 proteins
CC will be useful for a better understanding of cell growth, malignancy,
CC signal transduction processes, etc. New candidate drugs can be
CC identified, and their specificities (e.g. pharmacological activities) can
CC be assessed using the method of the invention

SQ Sequence 2873 BP; 903 A; 506 C; 623 G; 841 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	7.79e-28	Length:	2873
Score:	266.00	Matches:	44
Percent Similarity:	91.38%	Conservative:	9
Best Local Similarity:	75.86%	Mismatches:	5
Query Match:	82.35%	Indels:	0
DB:	2	Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) x NAT39799 (1-2873)

QY 2 ValCYsgInValIleGlyMetTYRAspTYRThrAlaGlnAsnAspAspGluLeuAlaPhe 21

Db 1338 GTATGTCAGGAGTGCATATGATGACATATGACACCAATATGAAGATGAGCTCAGTTTC 1397

QY 22 AsnLYsgLYGlnIleIleAsnValLeuAsnLYsgLYGlnAspProAspTYRThrLYsgLYGlu 41

Db 1398 TCCAAAGGACACATCATTAATGATATGAAACAAGATGATCTGATGATGGCAAGAGAG 1457

QY 42 ValAsnGlyGlnValIleGlyLeuPheProSerAsnTYRValLYsLeuThrThrasp 59

Db 1458 ATCAAAGGGGTGACTGCTCTTCTTCTTCAACATACGTTAAGATGACACAGAGAC 1511

RESULT 15

ID ABK43586 standard; cDNA; 3746 BP.

XX ABK43586;

DT 05-JUN-2002 (first entry)

XX DNA encoding novel central nervous system protein #166.

XX Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasia; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminization;
KW endocrine disorder; diabetes; dancer; leukemia; neovascularization;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy; gene; ss.

XX Homo sapiens.

XX WO200155318-A2.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001MO-US001332.

XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214866P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0233423P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0234984P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SW;

WPI; 2001-581633/65.

P-PSDB; AAU87256.

PT New isolated nucleic acid encoding a protein for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives.

Claim 1; SEQ ID NO 176; 837bp; English.

XX The invention describes an isolated nucleic acid molecule (1) encoding a

CC novel central nervous system protein. (1) and polypeptides (III) encoded
CC by (I), are used to treat a medical conditions and in diagnosis of a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphasia,
CC adenocarcinomas and irritable bowel syndrome, reproductive system
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,

Alignment Scores:

Pred. No.:	1.09e-27	Length:	3746
Score:	266.00	Matches:	44
Percent Similarity:	91.38%	Conservative:	9
Best Local Similarity:	75.86%	Mismatches:	5
Query Match:	82.35%	Indels:	0
DB:	4	Gaps:	0

US-09-720-934-2_COPY_1080_1138 (1-59) x ABR43586 (1-3746)

QY	2	ValCysGlnValIleGlyMetTyrAspTyrThrAlaGlnAspAspGluLeuAlaPhe	21
DB	1264	GTAATGCAAGTGTATTCCTATGTAATGCAAGCAAAATATATGAATGAGCTCAGTTTC	1323
QY	22	AsnLysGlyGlnIleIleAsnValLeuAsnLysGluAspProAspTyrPheLysGlyGlu	41
DB	1324	TCCAAAGGACAACTCATATGTTATGAACTAAGATGCTTATTTGGTGCAGAGAGG	1383
QY	42	ValAsnGlyGlnValGlyLeuPheProSerAsnTyrValLysLeuThrThraSp	59
DB	1384	ATCAACGGGGTGACTGGTCTTCTTCTCAACTAGCTTAAGATGACAGACAGC	1437

Search completed: July 1, 2004, 13:32:11
Job time : 189.112 secs

PA (CYTO-) CYTOGEN CORP

PA (UYN-) UNIV NORTH CAROLINA.
 XX Sparks AB, Hoffman N, Kay BK, Fowlkes DM, Mcconnell SJ;
 XX WPI; 1996-465045/46.
 XX P-PSDB; AAM05395.
 PT Identifying polypeptide(s) having specific functional domain (esp. SH3
 PT domain) - comprises detecting selective binding to recognition unit,
 PT regardless of sequence homology
 PS Claim 76, Fig 50, 174pp; English.
 XX AAT39786-T39803 represent novel human and mouse genes encoding Src-
 CC homology region 3 (SH3) domain containing proteins that can be used in
 CC the method of the invention. SH3 domain containing proteins play a role
 CC in signalling and structural elements of cells. The method of the
 CC invention is for identifying polypeptides containing functional domains
 CC of interest (especially SH3 domains). The method comprises contacting a
 CC multivalent recognition unit (RU) complex with a number of peptides and
 CC identifying polypeptides having a selective binding affinity for the RU
 CC complex. The method is based on functional similarities and does not rely
 CC on sequence similarities. Prior methods only gave limited success for
 CC identifying proteins which contain an SH3 domain due to the minimal
 CC sequence homology among known SH3 proteins. It has been found that small
 CC peptide RUs in multivalent form have reduced specificity for a given
 CC functional domain compared to monomer RUs. Multivalent RU complexes are
 CC particularly suited to screening for polypeptides containing functional
 CC domains that are similar to, but not identical in sequence to, the
 CC original target functional domain. The new method enables proteins having
 CC a common function to be identified. Identification of novel SH3 proteins
 CC will be useful for a better understanding of cell growth, malignancy,
 CC signal transduction processes, etc. New candidate drugs can be
 CC identified, and their specificities (e.g. pharmacological activities) can
 CC be assessed using the method of the invention
 XX Sequence 1389 BP; 404 A; 370 C; 357 G; 258 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 5.38e-33 Length: 1389
 Score: 341.00 Matches: 64
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_999_1062 (1-64) x AAT39795 (1-1389)
 QY 1 GlyGluGluIleAlaGlnValIleAlaSerYrThrAlaThrGlyProGluGlnLeuThr 20
 DB 949 GGAGAGAAATGCCAGGTATTGGCTCATACACCGCCACCGCCCGAGAGAGCTCACT 1008
 QY 21 LeuAlaProGlyGlyIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 40
 DB 1009 CTCGCCCTCGTGCAGCTGATTTCATCCAAAAAGAACCCAGTGGATGGTGGAGAGA 1068
 QY 41 GluLeuGlnAlaArgGlyIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 60
 DB 1069 GAGCTGCAAGCAGCGTGGAAAAAGCGCCAGATAGCTGCTCCAGCTAATTATGTAAG 1128
 QY 61 LeuLeuSerPro 64
 DB 1129 CTTCTAAGCCCT 1140
 RESULT 2
 AA234574
 ID AA234574 standard; cDNA; 3231 BP.
 XX
 AC AA234574;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human SH3D1A cDNA clone 9.

XX SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 XX Homo sapiens.
 OS
 XX Key
 FH Location/Qualifiers
 FT 2.1927
 FT CDS
 FT /*tag= a
 XX MO9953062-A2.
 XX 21-OCT-1999.
 XX 16-APR-1999; 99MO-US008371.
 XX 16-APR-1998; 98US-0082007P.
 XX (CEDA-) CEDAS SINAI HEALTH SYSTEM.
 PA Korenberg JR, Chen X;
 XX WPI; 1999-633829/54.
 DR P-PSDB; AAY32158.
 XX Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 PT Claim 2; Fig 14; 99pp; English.
 PS This is the nucleotide sequence of a non-full-length cDNA (clone 9)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukaemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC platelet disorder on chromosome 21, low platelets in deletion for 21,
 CC association of gains in chromosome 21 with leukaemias, neural
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukaemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 XX
 SQ Sequence 3231 BP; 1004 A; 721 C; 712 G; 794 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.54e-32 Length: 3231
 Score: 341.00 Matches: 64
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_999_1062 (1-64) x AA234574 (1-3231)
 QY 1 GlyGluGluIleAlaGlnValIleAlaSerYrThrAlaThrGlyProGluGlnLeuThr 20
 DB 1487 GGAGAGAAATGCCAGGTATTGGCTCATACACCGCCACCGCCCGAGAGAGCTCACT 1546

Qy	21	LeuAlaProGlyGlnLeuIleLeuIleAlaGlyLysAsnProGlyGlyTrrPrrGluGly	40
Db	1547	CTGGCCCTTGTCGTGATTTTGGATCCGAAAAAGAACCCAGGTGGATGTGGAGAGA	1666
Qy	41	GluLeuGlnAlaArgGlyLysLysArgGlnIleGlyTrrPhaProAlaSerValLys	60
Db	1607	GAGCTGAAACACGTGGAAAAAGCCCGATGAGCTGGTCCAGCTAATTATGTAAAG	1666
Qy	61	LeuLeuSerPro 64	
Db	1667	CTTCTAAGCCCT 1678	
RESULT 3			
ID	AAZ34570	standard; cDNA; 5199 BP.	
XX			
AC	AAZ34570;		
XX			
DT	01-FEB-2000	(first entry)	
XX			
DE	Human SH3D1A cDNA.		
XX			
KW	SH3D1A gene; human; Down's syndrome; leukaemia; cancer;		
KW	megakaryocytic abnormality; myeloproliferative disorder;		
KW	platelet disorder; neural disorder; thrombocytopenia;		
KW	haematopoietic disorder; cognitive dysfunction; microcephaly;		
KW	lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	208..3642	
FT		/*tag= a	
XX			
PN	WO9953062-A2.		
XX			
PD	21-OCT-1999.		
XX			
PF	16-APR-1999; 99WO-US0008371.		
XX			
PR	16-APR-1998; 98US-0082007P.		
XX			
PA	(CEDA-) CEDARS SINAI HEALTH SYSTEM.		
XX			
PI	Korenberg JR, Chen X;		
XX			
DR	WP1; 1999-633829/54.		
DR	P-PSDB; AAY32154.		
XX			
XX			
ES	Nucleic acid from the human SH3D1A gene and its products, useful for the		
XX	diagnosis and treatment of myeloproliferative disorders and leukemia.		
XX	Claim 2; Fig 5; 99pb; English.		

This is the nucleotide sequence of full-length cDNA corresponding to a novel human SH3 gene, termed the SH3D1A gene, that contributes to the development of platelets and the pathogenesis of leukaemias, both in general and in particular those involving the megakaryocytic lineage. The SH3D1A gene maps to the small candidate region for low platelets on chromosome 21. Sequencing of 5 different sizes of cDNA clone (see AA34570-74) suggests that at least 3 isoforms exist. The invention provides methods for the diagnosis and treatment of megakaryocytic abnormalities, myeloproliferative disorder, platelet disorder on leukaemia, neural disorders, thrombocytopenia, platelet disorder on chromosome 21, low platelets in deletion for 21, association of gains in chromosome 21 with leukaemias, neural abnormalities, dysfunctions and disorders including brain malformations and corresponding cognitive dysfunctions, microcephaly, lissencephaly, and colpocephaly. Methods are also provided for: suppressing cells unable to regulate themselves; screening for a somatic alteration in the SH3D1A gene; monitoring the progress and adequacy of a treatment; monitoring tumour risk progress or megakaryocytic abnormality, myeloproliferative disorder, haematopoietic disorder, platelet disorder or leukaemia; and treatment of a subject

CC	(including a prenatal subject) having megakaryocytic abnormality, CC myelodysplastic disorder, platelet disorder, leukaemia or neural CC disorder using a nucleic acid that expresses SH3D1A or its antisense CC nucleic acid
XX	
SQ	Sequence 5199 BP; 1615 A; 1152 C; 1214 G; 1218 T; 0 U; 0 Other;
Alignment Scores:	
Alignment Scores:	
Pred. No.:	2.8e-32
Score:	341.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	2
	Gaps: 0
US-09-720-934-2_COPY_999_1062 (1-64) x AA234570 (1-5199)	
QY	1 GYGLUGLUILEALAGINVAILEALASERTYTRRAIATHRGILPRGILGILNEUTHR 20
Db	3202 GGAAGAAATTCGCCAGATTTCCTCATACACCGCACCGGCCGAGAGTCACT 3265
QY	21 LEUALPRGILGLINLEUILEUILEARGYSLSAASPRGILGYLYTRTPRGLUGLY 40
Db	3262 CTCGCCCTCGCTCAGCTGATTTTGATCCGAAAAAGAACCCAGGTGATGTCGAGCA 3321
QY	41 GLULEUGLALAIRGILYLYSLYSARGGILNIEGLYTRPRNEPRCALASNTYRYALLYS 60
Db	3322 GACCTGCAGCACGTCGGAAAAAGCCACAGATAGGCTGTCCTCCACGATTAATTATGTAAG 3381
QY	61 LEULEUSERPRO 64
Db	3382 CTTCTAAGCCCT 3393
RESULT 4	
ID	AA234572
XX	AA234572 standard; cDNA; 5195 BP.
AC	AA234572;
DT	01-FEB-2000 (first entry)
XX	
DE	Human SH3D1A cDNA clone 11.
XX	
KW	SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
KW	megakaryocytic abnormality; myelodysplastic disorder;
KW	platelet disorder; neural disorder; thrombocytopenia;
KW	haematopoietic disorder; cognitive dysfunction; microcephaly;
KW	lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
OS	Homo sapiens.
XX	
XX	
Key	Location/Qualifiers
FT	239.. 3886
FT	/*tag= a
XX	
PN	W09953062-A2.
XX	
PD	21-OCT-1999.
XX	
PF	16-APR-1999; 99WO-US008371.
XX	
PR	16-APR-1998; 98US-0082007P.
XX	
FA	(CEDA-) CEDARS SINAI HEALTH SYSTEM.
XX	
PI	Korenberg JR, Chen X;
XX	
DR	WPI: 1999-633829/54.
DR	P-FSDB; AAY32156.
XX	
XX	
PT	Nucleic acid from the human SH3D1A gene and its products, useful for the
PT	diagnosis and treatment of myelodysplastic disorders and leukemia.

PS Claim 2; Fig 10; 99pp; English.

XX This is the nucleotide sequence of full-length cDNA (clone 11)

CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that

CC contributes to the development of platelets and the pathogenesis of

CC leukemias, both in general and in particular those involving the

CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate

CC region for low platelets on chromosome 21. Sequencing of 5 different

CC sizes of cDNA clones from foetal brain (see AA234570-74) suggests that at

CC least 3 isoforms exist. The invention provides methods for the diagnosis

CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,

CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,

CC platelet disorder on chromosome 21, low platelets in deletion for 21,

CC association of gains in chromosome 21 with leukemias, neural

CC abnormalities, dysfunctions and disorders including brain malformations

CC and colpocephaly. Methods are also provided for: suppressing cells unable

CC to regulate themselves; screening for a somatic alteration in the SH3D1A

CC gene; monitoring the progress and adequacy of a treatment; monitoring

CC tumour risk progress or megakaryocytic abnormality, myeloproliferative

CC disorder, haematopoietic disorder, platelet disorder or leukaemia; and

CC treatment of a subject (including a prenatal subject) having

CC megakaryocytic abnormality, myeloproliferative disorder, platelet

CC disorder, leukaemia or neural disorder using a nucleic acid that

CC expresses SH3D1A or its antisense nucleic acid

XX

SQ Sequence 5195 BP; 1592 A; 1148 C; 1237 G; 1218 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6.69e-31 Length: 5195

Score: 330.00 Matches: 62

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 96.77% Indels: 0

DB: 2 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x AA234572 (1-5195)

QY 3 GtlltleaGlnValIleAlaSerTyrThrAlaThnGlyProGluGlnLeuThleuAla 22

Db 3452 GAATTGCCAGGTTATTGCTCATACACCGCACCGGCCCGAGACACTCTCCGC 3511

QY 23 ProGlyGlnLeuIleLeuIleArgLysLysAsnProGlyGlyTyrPTrpGluGluLeu 42

Db 3512 CTTGGTCAGCTGATTGATTCGAAAAAGAACCCAGGTGGATGGTGGAGAGAGCTG 3571

QY 43 GlnAlaArgGlyLysLysArgGlnIleGlyTyrPheProAlaAsnTyrValLysLeuLeu 62

Db 3572 CAAGCAGCGGGAAAAAGCGGAGTAAGCTGTTCCAGCTAATTATGTAAGCTTCTA 3631

QY 63 SerPro 64

Db 3632 AGCCCT 3637

RESULT 5

AA234571

ID AA234571 standard; cDNA; 5458 BP.

XX AA234571;

AC AA234571;

DT 01-FEB-2000 (first entry)

XX

XX Human SH3D1A cDNA clone 21.

DE

XX SH3D1A gene; human; Down's syndrome; leukaemia; cancer;

KM megakaryocytic abnormality; myeloproliferative disorder;

KM platelet disorder; neural disorder; thrombocytopenia;

KM haematopoietic disorder; cognitive dysfunction; microcephaly;

KM lissencephaly; colpocephaly; helioproensephaly; diagnosis; therapy; ss.

XX

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT CDS 267..3929

XX /*tag= a

XX PN MO9953062-A2.

XX PD 21-OCT-1999.

XX

XX 16-APR-1999; 99MO-US008371.

XX

XX 16-APR-1998; 98US-0082007P.

XX

XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.

XX

XX PI Korenberg JR, Chen X;

XX

XX WPI; 1999-633829/54.

XX

XX P-PSDB; AAY32155.

XX

XX Nucleic acid from the human SH3D1A gene and its products, useful for the

XX diagnosis and treatment of myeloproliferative disorders and leukemia.

XX

XX Claim 2; Fig 8; 99pp; English.

XX

XX This is the nucleotide sequence of full-length cDNA (clone 21)

XX corresponding to a novel human SH3 gene, termed the SH3D1A gene, that

XX contributes to the development of platelets and the pathogenesis of

XX leukemias, both in general and in particular those involving the

XX megakaryocytic lineage. The SH3D1A gene maps to the small candidate

XX region for low platelets on chromosome 21. Sequencing of 5 different

XX sizes of cDNA clones from foetal brain (see AA234570-74) suggests that at

XX least 3 isoforms exist. The invention provides methods for the diagnosis

XX and treatment of megakaryocytic abnormality, myeloproliferative disorder,

XX platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,

XX platelet disorder on chromosome 21, low platelets in deletion for 21,

XX association of gains in chromosome 21 with leukemias, neural

XX abnormalities, dysfunctions and disorders including brain malformations

XX and colpocephaly. Methods are also provided for: suppressing cells unable

XX to regulate themselves; screening for a somatic alteration in the SH3D1A

XX gene; monitoring the progress and adequacy of a treatment; monitoring

XX tumour risk progress or megakaryocytic abnormality, myeloproliferative

XX disorder, haematopoietic disorder, platelet disorder or leukaemia; and

XX treatment of a subject (including a prenatal subject) having

XX megakaryocytic abnormality, myeloproliferative disorder, platelet

XX disorder, leukaemia or neural disorder using a nucleic acid that

XX expresses SH3D1A or its antisense nucleic acid

XX

SQ Sequence 5458 BP; 1671 A; 1201 C; 1307 G; 1279 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7.12e-31 Length: 5458

Score: 330.00 Matches: 62

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 96.77% Indels: 0

DB: 2 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x AA234571 (1-5458)

QY 3 GtlltleaGlnValIleAlaSerTyrThrAlaThnGlyProGluGlnLeuThleuAla 22

Db 3495 GAATTGCCAGGTTATTGCTCATACACCGCACCGGCCCGAGACACTCTCCGC 3554

QY 23 ProGlyGlnLeuIleLeuIleArgLysLysAsnProGlyGlyTyrPTrpGluGluLeu 42

Db 3555 CTTGGTCAGCTGATTGATTCGAAAAAGAACCCAGGTGGATGGTGGAGAGAGCTG 3614

QY 43 GlnAlaArgGlyLysLysArgGlnIleGlyTyrPheProAlaAsnTyrValLysLeuLeu 62

Db 3615 CAAGCAGCGGGAAAAAGCGGAGTAAGCTGTTCCAGCTAATTATGTAAGCTTCTA 3674

QY 63 SerPro 64


```

Db          3675 AGCCT 3680

RESULT 6
AAZ39009
ID  AAZ39009 standard; cDNA; 3723 BP.
XX
XX
AC  AAZ39009;
XX
DT  28-FEB-2000 (first entry)
XX
DE  Mouse Ese1 coding sequence.
XX
KM  Mouse; murine; Ese1, Ese2; endocytosis; vesicular trafficking;
KM  regulation; actin cytoskeleton; detection; cancer; infection;
KM  EH-domain and SH3-domain regulator of endocytosis; anticancer;
XX  antiproliferative; antiviral; ss.
XX
OS  Mus sp.
XX
PN  W09955728-A2.
XX
PD  04-NOV-1999.
XX
PF  27-APR-1999; 99WC-CA000375.
XX
PR  27-APR-1998; 98CA-02230201.
XX  05-FEB-1999; 99US-0118739P.
PA  (HSCR-) HSC RES & DEV LP.
XX
PI  Egan SE, Wang W, Sengar A;
XX
DR  WPI; 2000-052802/04.
XX  F-PSDB; AA57444.
XX
PT  New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
PT  endocytosis, used e.g. for treating cancer or preventing viral infection.
XX
PS  Claim 6; Page 40-42; 99p; English.

The present sequence encodes mouse Ese1. The present invention
specifically describes mammalian Ese1 and 2 proteins (1) and their splice
variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (1)
are involved in regulation of clathrin-mediated endocytosis (as a complex
with Eps15 protein), vesicular trafficking and actin cytoskeleton.
Generally (1) (or its (ant)agonists, mimetics, fragments and inactive
mutants): (1)-specific antibodies (Ab); sequences antisense to the (1)
polynucleotide; agents that downregulate expression of Ese genes or
antagonists of an Ese binding partner are used to treat diseases
associated with undesirable endocytosis and resulting changes in cellular
function. Particularly overexpression of Ese1 is used to block clathrin-
mediated endocytosis in vivo or in cell cultures, while administration of
(1) or Ab are used to suppress abnormal proliferation of cells that can
be stimulated to proliferate by a growth factor receptor; and similar
compounds (also inactive Ese mutants) can be used to prevent viral
infection. Endocytosis may also be regulated, in vivo or in cell
cultures, by forming an Ese-Eps15 complex, then binding dynamn to the
complex. Generally conditions that can be treated include cancer;
abnormal cell division or migration; viral infection; or abnormal
receptor signalling, tissue development or synaptic transmission

SQ  Sequence 3723 BP; 1065 A; 941 C; 1025 G; 692 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,87e-30 Length: 3723
Score: 325.00 Matches: 61
Percent Similarity: 98.39% Conservative: 0
Best Local Similarity: 98.39% Mismatches: 1
Query Match: 95.31% Indels: 0
DB: 3 Gaps: 0

US-09-720-934-2_COPY_999_1062 (1-64) x AAZ39009 (1-3723)

```


PA (HSCR-) HSC RES & DEV LP.
 XX Egan SE, Wang W, Sengar A;
 PI WPI: 2000-052802/04.
 DR P-PSDB; MAY57449.
 XX
 PT New nucleic acid encoding Esei and 2 proteins, involved in regulation of
 PT endocytosis, used e.g. for treating cancer or preventing viral infection.
 XX
 PS Claim 6; Page 56-59; 99pp; English.
 XX
 CC The present invention specifically describes mammalian Esei and 2
 CC proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
 CC regulator of endocytosis). (I) are involved in regulation of clathrin-
 CC mediated endocytosis (as a complex with Eps15 protein), vesicular
 CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
 CC mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
 CC sequences antisense to the (I) polynucleotide; agents that downregulate
 CC expression of Esei genes or antagonists of an Esei binding partner are used
 CC to treat diseases associated with undesirable endocytosis and resulting
 CC changes in cellular function. Particularly overexpression of Esei is used
 CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
 CC administration of (I) or Ab are used to suppress abnormal proliferation of
 CC cells that can be stimulated to proliferate by a growth factor receptor;
 CC and similar compounds (also inactive Esei mutants) can be used to prevent
 CC viral infection. Endocytosis may also be regulated, in vivo or in cell
 CC cultures, by forming an Esei-Eps15 complex, then binding dynamn to the
 CC complex. Generally conditions that can be treated include cancer;
 CC abnormal cell division or migration; viral infection; or abnormal
 CC receptor signalling, tissue development or synaptic transmission. The
 CC present sequence represents mouse Esei cDNA sequence
 XX
 SQ Sequence 5738 BP; 1556 A; 1456 C; 1613 G; 1113 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 3, 21e-30 Length: 5738
 Score: 325.00 Matches: 61
 Percent Similarity: 98.39% Conservative: 0
 Best Local Similarity: 98.39% Mismatches: 1
 Query Match: 95.31% Indels: 0
 DB: 3 Gaps: 0
 US-09-720-934-2_COPY_999_1062 (1-64) x AA239024 (1-5738)
 QY 3 GUILLEALAGVALLLEALSERLYRTHALATHGLYPROGLINLEUTHLEUALA 22
 DB 3467 GAATATGCCAGGTATGCTTCTTACGCTACTGATGCTCCGAAACACTCACTGGCT 3526
 QY 23 PROGLINLEULEULEULEALAGLYSLYASNPGLYGLYTRTPRPGIUGLYGUEN 42
 DB 3527 CCGGGCACACTGATTTGATCCGAAAAGAACCCAGGAGGAGTGGGAAGAGAACTG 3586
 QY 43 GINALAARGLYLYSLYASARGINILEGLYTRPPEPRALASNTYVALLYSLEULEU 62
 DB 3587 CAAGCTCGAGGAAAAGCGCCAGATAGGTTGTTCCAGCAATTATCTCAACTCTTA 3646
 QY 63 SerPro 64
 DB 3647 AGCCCC 3652
 RESULT 10
 ABQ55007 ID ABQ55007 standard; cDNA; 2067 BP.
 XX
 AC ABQ55007;
 XX
 DT 22-AUG-2002 (first entry)
 XX
 DE Human ovarian antigen HVBX28 cDNA, SEQ ID NO:887.
 XX
 KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;

KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; chromosome 21q22.1-22.2;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO20020677-A1.
 XX
 PD 03-JAN-2002.
 XX
 PF 07-JUN-2001; 2001MO-US018569.
 XX
 PR 07-JUN-2000; 2000US-0209467P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI: 2002-147878/19.
 DR P-PSDB; ABP41930.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
 PT cancer), immune disorders, cardiovascular disorders and neurological
 PT diseases.
 XX
 PS Claim 1; SEQ ID NO 887; 2922pp; English.
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2067 BP; 614 A; 426 C; 418 G; 605 T; 0 U; 4 Other;
 XX
 Alignment Scores:
 Pred. No.: 5, 07e-30 Length: 2067
 Score: 319.00 Matches: 59
 Percent Similarity: 98.39% Conservative: 2
 Best Local Similarity: 95.16% Mismatches: 1
 Query Match: 93.55% Indels: 0

XX Identifying polypeptide(s) having specific functional domain (esp. SH3 domain) - comprises detecting selective binding to recognition unit, regardless of sequence homology.

PS Claim 76; Fig 52; 174pp; English.

CC AAT39786-T39803 represent novel human and mouse genes encoding Src-homology region 3 (SH3) domain containing proteins that can be used in the method of the invention. SH3 domain containing proteins play a role in signalling and structural elements of cells. The method of the invention is for identifying polypeptides containing functional domains of interest (especially SH3 domains). The method comprises contacting a multivalent recognition unit (RU) complex with a number of peptides and identifying polypeptides having a selective binding affinity for the RU complex. The method is based on functional similarities and does not rely on sequence similarities. Prior methods only gave limited success for identifying proteins which contain an SH3 domain due to the minimal sequence homology among known SH3 proteins. It has been found that small peptide RUs in multivalent form have reduced specificity for a given functional domain compared to monomer RUs. Multivalent RU complexes are particularly suited to screening for polypeptides containing functional domains that are similar to, but not identical in sequence to, the original target functional domain. The new method enables proteins having a common function to be identified. Identification of novel SH3 proteins will be useful for a better understanding of cell growth, malignancy, signal transduction processes, etc. New candidate drugs can be identified, and their specificities (e.g. pharmacological activities) can be assessed using the method of the invention.

SO Sequence 747 BP; 244 A; 133 C; 186 G; 184 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,688-22	Length:	747
Score:	253.00	Matches:	47
Percent Similarity:	87.10%	Conservative:	7
Best Local Similarity:	75.81%	Mismatches:	8
Query Match:	74.19%	Indels:	0
DB:	2	Gaps:	0

US-09-720-934-2_COPY_999_1062 (1-64) x AAT39796 (1-747)

QY	3	GluilealaglnvalillealaserTyrThrAlathrglyProgluglnleuthleuAla	22
DB	334	GAGATTGCTCAGGTAACCTCAGCATATGTTGCTTCGTTGCAACAATTGCTTGA	393
QY	23	ProgluglnleuileuileuilearglylsasnpProglyTyrTrrpglulglnleu	42
DB	394	CCAGGACGTTAAATTAATTTCTAAAGAAAATACAGGCGTGGTGGCAAGAGTTA	453
QY	43	GlnAlaArgGlyLysLysArgGlnIleGlyTrrPheProAlaAsnTyVallylsleu	62
DB	454	CAGGCGCAAGGAAAAAGCGACGAAAGAGTGTTCCTGCCAGTCATGTTAAACTTTTG	513
QY	63	SerPro 64	
DB	514	GGTCCA 519	

RESULT 13
AAT39799
ID AAT39799 standard; DNA; 2873 BP.

XX AAT39799;

XX 19-FEB-1998 (first entry)

XX Human clone 65 gene.

XX Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;
XX cellular signalling element; cellular structural element; malignancy;
XX protein identification; functional domain; protein screening;
XX cellular signal transduction process; ss.

OS Homo sapiens.

XX Key	Location/Qualifiers
XX CDS	3..1532
XX FT	/*tag= a

XX W09631625-A1.

XX 10-OCT-1996.

XX 04-APR-1996; 96MO-US004454.

XX 07-APR-1995; 95US-00417872.

XX 03-APR-1996; 96US-00630915.

XX (CYTO-) CYTOGEN CORP.

XX (YUNC-) UNIV NORTH CAROLINA.

XX Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;

XX WPI; 1996-465045/46.

XX P-PDB; AAM05399.

XX Claim 76; Fig 58; 174pp; English.

CC AAT39786-T39803 represent novel human and mouse genes encoding Src-homology region 3 (SH3) domain containing proteins that can be used in the method of the invention. SH3 domain containing proteins play a role in signalling and structural elements of cells. The method of the invention is for identifying polypeptides containing functional domains of interest (especially SH3 domains). The method comprises contacting a multivalent recognition unit (RU) complex with a number of peptides and identifying polypeptides having a selective binding affinity for the RU complex. The method is based on functional similarities and does not rely on sequence similarities. Prior methods only gave limited success for identifying proteins which contain an SH3 domain due to the minimal sequence homology among known SH3 proteins. It has been found that small peptide RUs in multivalent form have reduced specificity for a given functional domain compared to monomer RUs. Multivalent RU complexes are particularly suited to screening for polypeptides containing functional domains that are similar to, but not identical in sequence to, the original target functional domain. The new method enables proteins having a common function to be identified. Identification of novel SH3 proteins will be useful for a better understanding of cell growth, malignancy, signal transduction processes, etc. New candidate drugs can be identified, and their specificities (e.g. pharmacological activities) can be assessed using the method of the invention.

SO Sequence 2873 BP; 903 A; 506 C; 623 G; 841 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,448-21	Length:	2873
Score:	253.00	Matches:	47
Percent Similarity:	87.10%	Conservative:	7
Best Local Similarity:	75.81%	Mismatches:	8
Query Match:	74.19%	Indels:	0
DB:	2	Gaps:	0

US-09-720-934-2_COPY_999_1062 (1-64) x AAT39799 (1-2873)

QY	3	GluilealaglnvalillealaserTyrThrAlathrglyProgluglnleuthleuAla	22
DB	1119	GAGATTGCTCAGGTAACCTCAGCATATGTTGCTTCGTTGCAACAATTGCTTGA	1178

QY	23	ProgluglnleuileuileuilearglylsasnpProglyTyrTrrpglulglnleu	42
DB	1179	CCAGGACGTTAAATTAATTTCTAAAGAAAATACAGGCGTGGTGGCAAGAGTTA	1238

QY	43	GlnAlaArgGlyLysLysArgGlnIleGlyTrrPheProAlaAsnTyVallylsleu	62
----	----	---	----

Db 1239 CAGGCCAGAGGAAAAAGGACGAAAGGATGTTCTCTCCAGCATGTTAACTTTTG 1298
 QY 63 SerPro 64
 Db 1299 GGTCCA 1304
 RESULT 14
 AA239011
 ID AA239011 standard; cDNA; 3593 BP.
 AC AA239011;
 XX 28-FEB-2000 (first entry)
 DT Mouse E2e2 coding sequence.
 DE
 XX Mouse; murine; E2e2; endocytosis; vesicular trafficking;
 KW regulation; actin cytoskeleton; detection; cancer; infection;
 KW BH-domain and SH3-domain regulator of endocytosis; anticancer;
 KW antiproliferative; antiviral; s.
 XX
 OS Mus sp.
 PN WO955728-A2.
 XX 04-NOV-1999.
 PD
 XX 27-APR-1999; 99WO-CA000375.
 PF
 XX 27-APR-1998; 98CA-02230201.
 PR 05-FEB-1999; 99US-0118739P.
 XX
 PA (HSCR-) HSC RES & DEV LP.
 XX Egan SE, Wang W, Sengar A;
 XX WPI; 2000-052802/04.
 DR P-PSDB; AAY57445.
 DR
 XX New nucleic acid encoding E2e2 and 2 proteins, involved in regulation of
 PT endocytosis, used e.g. for treating cancer or preventing viral infection.
 FT
 XX
 PS Claim 25; Page 46-48; 99pp; English.
 XX
 CC The present sequence encodes mouse E2e2. The present invention
 CC specifically describes mammalian E2e2 and 2 proteins (I) and their splice
 CC variants (Bse = EH-domain and SH3-domain regulator of endocytosis). (I)
 CC are involved in regulation of clathrin-mediated endocytosis (as a complex
 CC with E2e2 protein), vesicular trafficking and actin cytoskeleton.
 CC Generally (I) (or its (ant)agonists, mimetics, fragments and inactive
 CC mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)
 CC polynucleotide; agents that downregulate expression of E2e2 genes or
 CC antagonists of an E2e2 binding partner are used to treat diseases
 CC associated with undesirable endocytosis and resulting changes in cellular
 CC function. Particularly overexpression of E2e2 is used to block clathrin-
 CC mediated endocytosis in vivo or in cell cultures, while administration of
 CC (I) is used to promote endocytosis of selected cells. (Ant)agonists of
 CC (I) or Ab are used to suppress abnormal proliferation of cells that can
 CC be stimulated to proliferate by a growth factor receptor; and similar
 CC compounds (also inactive E2e2 mutants) can be used to prevent viral
 CC infection. Endocytosis may also be regulated, in vivo or in cell
 CC cultures, by forming an E2e2-E2e2 complex, then binding dynamin to the
 CC complex. Generally conditions that can be treated include cancer;
 CC abnormal cell division or migration; viral infection; or abnormal
 CC receptor signalling, tissue development or synaptic transmission
 CC
 SQ Sequence 3593 BP; 1171 A; 783 C; 850 G; 789 T; 0 U; 0 Other;

Best Local Similarity: 75.81% Mismatches: 8
 Query Match: 74.19% Indels: 0
 DB: 3 Gaps: 0
 US-09-720-934-2_COPY_999_1062 (1-64) x AA239011 (1-3593)
 QY 3 GlnleAlaGlnValIleAlaSerYThrAlaThrGlyProGlnGlnLeuThrLeuAla 22
 Db 3046 GAGATCGCTCAAGTAACCTCAGCATATGCTGCTTCAGGACCTGACAGCTCAGCCTTGCG 3105
 QY 23 ProGlyGlnLeuIleLeuIleArglyLysAsnProGlyGlyTyrPrlGluGlyLeu 42
 Db 3106 CCGAGCAGCTAATATTAATCTTAAGAAAAACAACCGCGTGGTGGCAAGAGACTA 3165
 QY 43 GlnAlaArgGlyLysLysAsnArgGlnIleGlyTyrPrlPheProAlaAsnTyrValLysLeu 62
 Db 3166 CAGGCCAGAGGAGAAAGAACACAGACGAAAGGATGTTCTCTCCAGCATGTTAAAGCTGCTA 3225
 QY 63 SerPro 64
 Db 3226 GGTCCA 3231
 RESULT 15
 ABK43586
 ID ABK43586 standard; cDNA; 3746 BP.
 XX
 AC ABK43586;
 XX 05-JUN-2002 (first entry)
 DT
 XX
 DE DNA encoding novel central nervous system protein #166.
 XX
 KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
 KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
 KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
 KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
 KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
 KW adenocarcinoma; reproductive system disorder; testicular feminisation;
 KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
 KW respiratory disorder; renal disorder; kidney failure; blood disorder;
 KW myocardial infarction; wound healing; cell proliferation; skin aging;
 KW food additive; food preservative; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 PN WO200155318-A2.
 XX
 XX 02-AUG-2001.
 PD
 XX 17-JAN-2001; 2001WO-US001332.
 PF
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205151P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216476P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 11-JUL-2000; 2000US-0218290P.
 PR 14-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.

Alignment Scores:
 Pred. No.: 1.9e-21
 Score: 253.00
 Percent Similarity: 87.10%

Length: 3593
 Matches: 47
 Conservative: 7

PR	08-NOV-2000;	2000US-0246552P.
PR	08-NOV-2000;	2000US-0246552P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246609P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	08-NOV-2000;	2000US-0246613P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249264P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249298P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251889P.
PR	11-DEC-2000;	2000US-0251909P.
PR	05-JAN-2001;	2001US-0254097P.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA, Barash SC, Ruben SW;	
DR	MPI: 2001-581633/65.	
DR	P-PSDB; AAU87256.	
XX		
PT	New isolated nucleic acid encoding a protein for diagnosing, preventing,	
PT	treating or ameliorating medical conditions and used as food additives or	
XX	preservatives.	
PS	Claim 1; SEQ ID NO 176; 837bp; English.	
CC	The invention describes an isolated nucleic acid molecule (I) encoding a	
CC	novel central nervous system protein. (II) and polypeptides (III) encoded	
CC	by (II), are used to treat a medical conditions and in diagnosis of a	
CC	pathological condition. Disorders which are diagnosed or treated include	
CC	autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative	
CC	disorders e.g. neoplasms of the breast or liver, cardiovascular disorders	
CC	e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,	
CC	angiodystrophic lateral sclerosis, infections caused by bacteria, viruses	
CC	e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders	
CC	e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,	
CC	adenocarcinomas and irritable bowel syndrome, reproductive system	
CC	disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes	
CC	and pituitary dwarfism, cancers and disorders at the cellular level e.g.	
CC	leukemia, disorders involving neovascularisation e.g. malignancies,	
CC	respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.	
CC	acute kidney failure and blood related disorders e.g. myocardial	
CC	infarction. The polypeptides can also be used to aid wound healing and	

CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
 CC maintain organs before transplantation, for supporting cell culture of
 CC primary tissues, to regenerate tissues and in chemotaxis. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities, fat content, lipid, protein,

Alignment Scores:

Pred. No.:	2e-21	Length:	3746
Score:	253.00	Matches:	47
Percent Similarity:	87.10%	Conservative:	7
Best Local Similarity:	75.81%	Mismatches:	8
Query Match:	74.19%	Indels:	0
DB:	4	Gaps:	0

US-09-720-934-2_COPY_999_1062 (1-64) x ABK43586 (1-3746)

QY	3	GIUIIEA	aglnvali	leal	as	TYR	Thra	l	Thrg	l	YP	Progl	ug	ln	leu	Th	leu	Ala	22	
DB	1045	GAGATTGCTCAGGTA	CTTCAGGTA	TATGTTGCTTCTGTA	CAACTTAGCCTTGA														1104	
QY	23	Progl	Ygl	ln	leu	l	leu	l	leu	l	leu	l	leu	l	leu	l	leu	l	42	
DB	1105	CCAGGACAGTTA	TATTAATTC	TAAGAAAA	ATACA	GTGGTGTG	GCACAG	AGAGTTA											1164	
QY	43	Gl	Ala	l	arg	G	l	Y	S	l	Y	S	A	r	G	l	n	l	le	62
DB	1165	CAGGCCAGAG	AAAAAGCGAC	AGAAAGATG	GTTCCTGC	AGTC	CAAGTTAA	CTTTTG											1224	
QY	63	Ser	Pro	64																
DB	1225	GGTCCA	1230																	

Search completed: July 1, 2004, 13:32:04
 Job time : 205.546 secs

	9	323	99.5	5738	3	AAT39024	AAT39024	Mouse	Eset
	10	315	96.0	7435	5	AAS84763	AAS84763	DNA	encoc
	11	280	85.4	395	5	AAS84759	AAS84759	DNA	encoc
	12	251	76.5	206	4	AAB61319	AAB61319	Human	get
	13	251	76.5	206	4	AAI46323	AAI46323	Probe	#11
	14	251	76.5	206	6	ABS39875	ABS39875	Human	get
	15	251	76.5	206	6	ABS14329	ABS14329	Human	get
	16	251	76.5	458	4	ABA53551	ABA53551	Human	foot
	17	251	76.5	458	4	AAI33177	AAI33177	Probe	#11
	18	251	76.5	458	4	ABS26852	ABS26852	Human	int
	19	251	76.5	458	6	ABS01804	ABS01804	Human	get
	20	251	76.5	2873	2	AAT39799	AAT39799	Human	CD1
	21	251	76.5	3746	4	ABK43586	ABK43586	DNA	encoc
	22	251	76.5	4210	4	AAS27090	AAS27090	DNA	encoc
	23	251	76.5	4210	4	ADB93268	ADB93268	CDN	encoc
	24	251	76.5	4447	4	AAS02055	AAS02055	DNA	encoc
	25	251	76.5	5828	6	AAAL47247	AAAL47247	Allergic	
	26	251	76.5	6103	4	AAKS2332	AAKS2332	Human	pol
	27	250	76.2	3593	3	AAZ39011	AAZ39011	Mouse	Eset
	28	250	76.2	4625	3	AAZ39027	AAZ39027	Mouse	Eset
	29	250	76.2	4975	3	AAZ39026	AAZ39026	Mouse	Eset
	30	250	76.2	6014	5	AAAS84762	AAAS84762	DNA	encoc
	31	244.5	74.5	2874	5	ABQ55007	ABQ55007	Human	ova
	32	134.5	41.0	2067	6	ABL01095	ABL01095	Drosophil	il1
C	33	130.5	39.8	3981	4	ABL01095	ABL01095	Drosophil	il1
C	34	130.5	39.8	7225	4	AAH15090	AAH15090	Human	cel
C	35	129.5	39.5	571	4	AAH15090	AAH15090	Human	cel
C	36	129.5	39.5	1634	4	AAZ29226	AAZ29226	Human	cel
	37	129.5	39.5	1639	3	ABL07359	ABL07359	Human	cdn
	38	129.5	39.5	2531	4	ADBB63501	ADBB63501	Human	cdn
	39	129.5	39.5	2539	9	ACCG1330	ACCG1330	Gene	sequ
	40	129.5	39.5	4661	4	ABLO7358	ABLO7358	Drosophil	il1
	41	125.5	38.3	1356	7	ADBO80380	ADBO80380	Human	MDD
	42	125.5	38.3	2406	8	ADD71183	ADD71183	Human	int
	43	125.5	38.3	2626	9	AAI32643	AAI32643	Human	CDN
	44	125.5	38.3	2801	4	AAZ27080	AAZ27080	CDNA	enco
C	45	125.5	38.3	2803	4	AAZ27080	AAZ27080	CDNA	enco

ALIGNMENTS

RESULT 1	ID	AAT39795 standard; DNA; 1389 BP.
AAT39795		
XX AC	AAT39795;	
XX DT	19-FEB-1998 (first entry)	
XX DE	Human SH3P17 gene.	
XX KW	Src-homology region 3 domain; human; mouse; SH3 domain; cell growth; cellular signalling element; cellular structural element; malignancy; protein identification; functional domain; protein screening;	
XX KM	cellular signal transduction process; ss.	
XX OS	Homo sapiens.	
XX FH	Key	Location/Qualifiers
XX FT	CDS	1..1389
XX FT		/tag= a
XX FT		/product= "SH3P17"
XX EN	W09631625-A1.	
XX PD	10-OCT-1996.	
XX PF	04-APR-1996;	96WO-US004454.
XX PR	07-APR-1995;	95US-00417872.
XX DR	03-APR-1996;	96US-00630915.
XX NA	(CYTO-) CYTOGEN CORP.	

PA (UYN-) UNIV NORTH CAROLINA.
 XX Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ,
 PI WPI; 1996-465045/46.
 DR P-PSDB; AAM05395.
 XX Identifying polypeptide(s) having specific functional domain (esp. SH3
 PT domain) - comprises detecting selective binding to recognition unit,
 PT regardless of sequence homology
 XX
 PS Claim 76; Fig 50; 174pp; English.
 XX AAT39786-T39803 represent novel human and mouse genes encoding Src-
 CC homology region 3 (SH3) domain containing proteins that can be used in
 CC the method of the invention. SH3 domain containing proteins play a role
 CC in signalling and structural elements of cells. The method of the
 CC invention is for identifying polypeptides containing functional domains
 CC of interest (especially SH3 domains). The method comprises contacting a
 CC multivalent recognition unit (RU) complex with a number of peptides and
 CC identifying polypeptides having a selective binding affinity for the RU
 CC complex. The method is based on functional similarities and does not rely
 CC on sequence similarities. Prior methods only gave limited success for
 CC identifying proteins which contain an SH3 domain due to the minimal
 CC sequence homology among known SH3 proteins. It has been found that small
 CC peptide RUs in multivalent form have reduced specificity for a given
 CC functional domain compared to monomer RUs. Multivalent RU complexes are
 CC particularly suited to screening for polypeptides containing functional
 CC domains that are similar to, but not identical in sequence to, the
 CC original target functional domain. The new method enables proteins having
 CC a common function to be identified. Identification of novel SH3 proteins
 CC will be useful for a better understanding of cell growth, malignancy,
 CC signal transduction processes, etc. New candidate drugs can be
 CC identified, and their specificities (e.g. pharmacological activities) can
 CC be assessed using the method of the invention
 XX
 SQ Sequence 1389 BP; 404 A; 370 C; 357 G; 258 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 6.04e-37 Length: 1389
 Score: 328.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_908_966 (1-59) x AAT39795 (1-1389)
 QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTPrArgAlaLysAspAsnHisIleu 20
 Db 676 GTGAGGGGGCTACAGCTCAAGCCCTATATCCTTGAGAGCCAAAGAACCAACCACTTA 735
 QY 21 AsnPhenAsnLysAsnAspValIleThrValLeuGluGlnGlnAspMetTyrTrpPheGly 40
 Db 736 AATTTTAACAATAATATGTATCTATCAACCGTCTGGAAACACAGACATGTGGTGTGA 795
 QY 41 GluValGlnGlyGlnLysGlyTyrPheProLysSerTyrValLysLeuIleSerGly 59
 Db 796 GAAGTTCAAGGTCTAGAGGGTGGTTCCCAAGCTTACGTGAACCACTTCATTTCAAGGG 852
 RESULT 2
 AA234574
 ID AA234574 standard; cDNA; 3231 BP.
 XX
 AC AA234574;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human SH3D1A cDNA clone 9.
 XX
 XX SH3D1A gene; human; Down's syndrome; leukemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;

KW haematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 2.1927
 FT CDS /*tag= a
 XX
 PN W09953062-A2.
 XX
 XX 21-OCT-1999.
 PD
 XX
 PE 16-APR-1999; 99MO-US008371.
 PF
 XX 16-APR-1998; 98US-0082007P.
 PR
 XX (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 PA
 XX Korenberg JR, Chen X;
 PI WPI; 1999-633829/54.
 DR P-PSDB; AAY32158.
 XX
 PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 PS Claim 2; Fig 14; 99pp; English.
 XX
 XX This is the nucleotide sequence of a non-full-length cDNA (clone 9)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC platelet disorder on chromosome 21, low platelets in deletion for 21,
 CC association of gains in chromosome 21 with leukemias, neural
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, haematopoietic disorder, platelet disorder or leukemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 XX
 SQ Sequence 3231 BP; 1004 A; 721 C; 712 G; 794 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.8e-36 Length: 3231
 Score: 328.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-720-934-2_COPY_908_966 (1-59) x AA234574 (1-3231)
 QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTPrArgAlaLysAspAsnHisIleu 20
 Db 1214 GTGAGGGGGCTACAGCTCAAGCCCTATATCCTTGAGAGCCAAAGAACCAACCACTTA 1273
 QY 21 AsnPhenAsnLysAsnAspValIleThrValLeuGluGlnGlnAspMetTyrTrpPheGly 40
 Db 1274 AATTTTAACAATAATATGTATCTATCAACCGTCTGGAAACACAGACATGTGGTGTGA 1333

QY 41 GluValGlnGlyGlnIleGlyTyrPheProIysSerTyrValIleuLeuSerGly 59
 Db 1334 GAAGTTCAGGCTCAGAGGCTGTGCTCCCAAGCTTACGTAACACTATTTCAGG 1390
 RESULT 3
 AA234572
 ID AA234572 standard; cDNA; 5195 BP.
 AC AA234572;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human SH3D1A cDNA clone 11.
 XX
 KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW hematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 239..3886
 FT /*tag= a
 XX
 PN WO9953062-A2.
 XX
 PD 21-OCT-1999.
 XX
 PF 16-APR-1999; 99WO-US008371.
 XX
 PR 16-APR-1998; 98US-0082007P.
 XX
 PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 XX
 PI Korenberg JR, Chen X;
 XX
 DR WPI; 1999-633829/54.
 DR P-PSDB; AAY32156.
 XX
 PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 FT diagnosis and treatment of myeloproliferative disorders and leukemia.
 XX
 PS Claim 2; Fig 10; 99p; English.
 XX
 CC This is the nucleotide sequence of full-length cDNA (clone 11)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukaemia, neural disorders, thrombocytopenia,
 CC association of gains in chromosome 21, low platelets in deletion for 21,
 CC abnormalities, dysfunctions and disorders including brain malformations
 CC and corresponding cognitive dysfunctions, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumour risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, hematopoietic disorder, platelet disorder or leukaemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukaemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid
 XX
 SQ Sequence 5195 BP; 1592 A; 1148 C; 1237 G; 1218 T; 0 U; 0 Other;
 Alignment Scores:

Pred. No.: 3,33e-36 Length: 5195
 Score: 328.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0
 US-09-720-934-2_COPY_908_966 (1-59) x AA234572 (1-5195)
 QY 1 ValGlnGlyLeuGlnAlaGlnAlaLeuTyrProTyrPheArgAlaIleuSAspAsnHisLeu 20
 Db 2960 GTGAGGGGGCTACAGGCTCAGGCCCTATTCCTTGGAGAGCCAAAAAGCAACACACTTA 3019
 QY 21 AsnPheAsnIleSAspValIleThrValIleGlnGlnIleSpmcTyrTrpPheGly 40
 Db 3020 AATTATTAACAAAATAGATGTCATCACCGCTCTGGAAAGCAAGACATGTGGTTTGA 3079
 QY 41 GluValGlnGlyGlnIleGlyTyrPheProIysSerTyrValIleuLeuSerGly 59
 Db 3080 GAAGTTCAGGCTCAGAGGCTGTGCTCCCAAGCTTACGTAACACTATTTCAGG 3136
 RESULT 4
 AA234570
 ID AA234570 standard; cDNA; 5199 BP.
 AC AA234570;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human SH3D1A cDNA.
 XX
 KW SH3D1A gene; human; Down's syndrome; leukaemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW hematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 208..3642
 FT /*tag= a
 XX
 PN WO9953062-A2.
 XX
 PD 21-OCT-1999.
 XX
 PF 16-APR-1999; 99WO-US008371.
 XX
 PR 16-APR-1998; 98US-0082007P.
 XX
 PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 XX
 PI Korenberg JR, Chen X;
 XX
 DR WPI; 1999-633829/54.
 DR P-PSDB; AAY32154.
 XX
 PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 FT diagnosis and treatment of myeloproliferative disorders and leukemia.
 XX
 PS Claim 2; Fig 5; 99p; English.
 XX
 CC This is the nucleotide sequence of full-length cDNA corresponding to a
 CC novel human SH3 gene, termed the SH3D1A gene, that contributes to the
 CC development of platelets and the pathogenesis of leukemias, both in
 CC general and in particular those involving the megakaryocytic lineage. The
 CC SH3D1A gene maps to the small candidate region for low platelets on
 CC chromosome 21. Sequencing of 5 different sizes of cDNA clone (see
 CC AA234570-74) suggests that at least 3 isoforms exist. The invention
 CC provides methods for the diagnosis and treatment of megakaryocytic
 CC abnormality, myeloproliferative disorder, platelet disorder, acute
 CC leukaemia, neural disorders, thrombocytopenia, platelet disorder on

CC chromosome 21, low platelets in deletion for 21, association of gains in
 CC chromosome 21 with leukemias, neural abnormalities, dysfunction and
 CC disorders including brain malformations and corresponding cognitive
 CC dysfunctions, microcephaly, lissencephaly, and colpocephaly. Methods are
 CC also provided for: suppressing cells unable to regulate themselves;
 CC screening for a somatic alteration in the SH3D1A gene; monitoring the
 CC progress and adequacy of a treatment; monitoring tumor risk progress or
 CC megakaryocytic abnormality, myeloproliferative disorder, hematopoietic
 CC disorder, platelet disorder or leukemia; and treatment of a subject
 CC (including a prenatal subject) having megakaryocytic abnormality,
 CC myeloproliferative disorder, platelet disorder, leukemia or neural
 CC disorder using a nucleic acid that expresses SH3D1A or its antisense
 CC nucleic acid

XX
 XX
 SQ Sequence 5199 BP; 1615 A; 1152 C; 1214 G; 1218 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3.34e-36 Length: 5199
 Score: 328.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x AA234570 (1-5199)

OY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyPProTPrArgAlaLysLysAspAsnHisLeu 20
 Db 2929 GTGAGGGGGCTACAAAGCTCAAGCCCTATATCTTGAGAGCCCAAAAAGACACCACTTA 2998

OY 21 AsnPheAsnLysAsnAspValIleThrValLeuGluGlnGlnAspMetTTrpPheGly 40
 Db 2989 AATTTTAACAAAAGATGATCATCAACCGCTCTGGAACGACAGCATGTGGTTGGA 3048

OY 41 GluValGlnGlyGlnLysGlyTTrpPheProLysSerTyValLysLeuIleSerGly 59
 Db 3049 GAAGTTCAAGGTCAAGAGGTTGGTCCCAAGCTTACGGAACCTATTTCAGGG 3105

RESULT 5
 AA234571
 ID AA234571 standard; cDNA; 5458 BP.

XX
 AC AA234571;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human SH3D1A cDNA clone 21.

XX
 KW SH3D1A gene; human; Down's syndrome; leukemia; cancer;
 KW megakaryocytic abnormality; myeloproliferative disorder;
 KW platelet disorder; neural disorder; thrombocytopenia;
 KW hematopoietic disorder; cognitive dysfunction; microcephaly;
 KW lissencephaly; colpocephaly; holoprosencephaly; diagnosis; therapy; ss.

XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 267..3929
 FT /*tag= a
 XX
 PN W09953062-A2.

XX
 PD 21-OCT-1999.
 XX
 PF 16-APR-1999; 99WO-US008371.
 XX
 PR 16-APR-1998; 98US-0082007P.
 XX
 PA (CEDA-) CEDARS SINAI HEALTH SYSTEM.
 XX
 PI Korenberg JR, Chen X;
 XX
 DR WPI, 1999-633829/54.

DR P-PSDB; AAY32155.
 XX
 PT Nucleic acid from the human SH3D1A gene and its products, useful for the
 PT diagnosis and treatment of myeloproliferative disorders and leukemia.
 XX
 PS Claim 2; Fig 8; 99p; English.

XX
 CC This is the nucleotide sequence of full-length cDNA (clone 21)
 CC corresponding to a novel human SH3 gene, termed the SH3D1A gene, that
 CC contributes to the development of platelets and the pathogenesis of
 CC leukemias, both in general and in particular those involving the
 CC megakaryocytic lineage. The SH3D1A gene maps to the small candidate
 CC region for low platelets on chromosome 21. Sequencing of 5 different
 CC sizes of cDNA clone from foetal brain (see AA234570-74) suggests that at
 CC least 3 isoforms exist. The invention provides methods for the diagnosis
 CC and treatment of megakaryocytic abnormality, myeloproliferative disorder,
 CC platelet disorder, acute leukemia, neural disorders, thrombocytopenia,
 CC platelet disorder on chromosome 21, low platelets in deletion for 21,
 CC association of gains in chromosome 21 with leukemias, neural
 CC abnormalities, dysfunction and disorders including brain malformations
 CC and corresponding cognitive dysfunction, microcephaly, lissencephaly,
 CC and colpocephaly. Methods are also provided for: suppressing cells unable
 CC to regulate themselves; screening for a somatic alteration in the SH3D1A
 CC gene; monitoring the progress and adequacy of a treatment; monitoring
 CC tumor risk progress or megakaryocytic abnormality, myeloproliferative
 CC disorder, hematopoietic disorder, platelet disorder or leukemia; and
 CC treatment of a subject (including a prenatal subject) having
 CC megakaryocytic abnormality, myeloproliferative disorder, platelet
 CC disorder, leukemia or neural disorder using a nucleic acid that
 CC expresses SH3D1A or its antisense nucleic acid

XX
 SQ Sequence 5458 BP; 1671 A; 1201 C; 1307 G; 1279 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3.55e-36 Length: 5458
 Score: 328.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x AA234571 (1-5458)

OY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyPProTPrArgAlaLysLysAspAsnHisLeu 20
 Db 3003 GTGAGGGGGCTACAAAGCTCAAGCCCTATATCTTGAGAGCCCAAAAAGACACCACTTA 3062

OY 21 AsnPheAsnLysAsnAspValIleThrValLeuGluGlnGlnAspMetTTrpPheGly 40
 Db 3063 AATTTTAACAAAAGATGATCATCAACCGCTCTGGAACGACAGCATGTGGTTGGA 3122

OY 41 GluValGlnGlyGlnLysGlyTTrpPheProLysSerTyValLysLeuIleSerGly 59
 Db 3123 GAAGTTCAAGGTCAAGAGGTTGGTCCCAAGCTTACGGAACCTATTTCAGGG 3179

RESULT 6
 AA239009
 ID AA239009 standard; cDNA; 3723 BP.

XX
 AC AA239009;
 XX
 DT 28-FEB-2000 (first entry)
 XX
 DE Mouse BseI coding sequence.
 XX
 KW Mouse; murine; BseI; Bse2; endocytosis; vesicular trafficking;
 KW regulation; actin cytoskeleton; detection; cancer; infection;
 KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
 KW antiproliferative; antiviral; ss.

XX
 OS Mus sp.
 XX
 PN W09955728-A2.

```
XX 04-NOV-1999.
PD
XX
XX 27-APR-1999; 99WO-CA000375.
PF
XX 27-APR-1998; 98CA-02230201.
PR 05-FEB-1999; 99US-0118739P.
PA (HSCR-) HSC RES & DEV LP.
XX
XX Egan SE, Wang W, Sengar A;
XX WPI, 2000-052802/04.
DR P-PSDB; AAY57444.
XX
XX New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.
PT
XX
PS Claim 6; Page 40-42; 99pp; English.
XX
XX The present sequence encodes mouse Ese1. The present invention
CC specifically describes mammalian Ese1 and 2 proteins (1) and their splice
CC variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (1)
CC are involved in regulation of clathrin-mediated endocytosis (as a complex
CC with Esp15 protein), vesicular trafficking and actin cytoskeleton.
CC Generally (1) (or its (ant)agonists, mimetics, fragments and inactive
CC mutants); (1)-specific antibodies (Ab); sequences antisense to the (1)
CC polynucleotide; agents that downregulate expression of Ese genes or
CC antagonists of an Ese binding partner are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of Ese1 is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (1) or Ab are used to promote endocytosis of selected cells. (Ant)agonists of
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-Esp15 complex, then binding dynamn to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission
XX
SQ Sequence 3723 BP; 1065 A; 941 C; 1025 G; 692 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,13e-35 Length: 3723
Score: 323.00 Matches: 58
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.31% Mismatches: 0
Query Match: 98.48% Indels: 0
DB: 3 Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x AA239009 (1-3723)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTTPaGAlaLysLysAspAsnHisLeu 20
DB 2716 GTGGAAGGGCTTCAAGCGCCCTGTATCCCTGGAGAGCCAAAAAAGACAACACCTTA 2775
QY 21 AsnPhaEnuLysAsnAspValIleThrValLeuGluGlnGlnAspMetTyrTTPheGly 40
DB 2776 AATTTTAAACAAGAGACGTATCAACGTTCTGGAACAGCAAGACATGTGGTGGT 2835
QY 41 GluValGlnGlyGlnLysGlyTyrPheProLysSerTyrValLysLeuIleSerGly 59
DB 2836 GAAGTTCAGATCGACAGAGGTTGTTCCCAAGTCTTACGTGAAGAACTCATTTCAAGG 2892
RESULT 7
AA239008
ID AA239008 standard, cDNA; 5082 BP.
XX
XX AA239008;
XX
XX 28-FEB-2000 (first entry)
DT
```

```
XX Mouse Ese1 full length cDNA sequence.
DE
XX
XX Mouse; murine; Ese1; endocytosis; vesicular trafficking;
KW regulation; actin cytoskeleton; detection; cancer; infection;
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;
KW antiproliferative; antiviral; ss.
XX
XX Mus sp.
XX
XX W09955728-A2.
XX
XX 04-NOV-1999.
XX
XX 27-APR-1999; 99WO-CA000375.
PF
XX 27-APR-1998; 98CA-02230201.
PR 05-FEB-1999; 99US-0118739P.
XX
XX (HSCR-) HSC RES & DEV LP.
XX
XX Egan SE, Wang W, Sengar A;
XX WPI, 2000-052802/04.
DR P-PSDB; AAY57444.
XX
XX New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
PT endocytosis, used e.g. for treating cancer or preventing viral infection.
PT
XX
PS Claim 6; Page 38-40; 99pp; English.
XX
XX The present sequence encodes mouse Ese1. The present invention
CC specifically describes mammalian Ese1 and 2 proteins (1) and their splice
CC variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (1)
CC are involved in regulation of clathrin-mediated endocytosis (as a complex
CC with Esp15 protein), vesicular trafficking and actin cytoskeleton.
CC Generally (1) (or its (ant)agonists, mimetics, fragments and inactive
CC mutants); (1)-specific antibodies (Ab); sequences antisense to the (1)
CC polynucleotide; agents that downregulate expression of Ese genes or
CC antagonists of an Ese binding partner are used to treat diseases
CC associated with undesirable endocytosis and resulting changes in cellular
CC function. Particularly overexpression of Ese1 is used to block clathrin-
CC mediated endocytosis in vivo or in cell cultures, while administration of
CC (1) or Ab are used to promote endocytosis of selected cells. (Ant)agonists of
CC be stimulated to proliferate by a growth factor receptor; and similar
CC compounds (also inactive Ese mutants) can be used to prevent viral
CC infection. Endocytosis may also be regulated, in vivo or in cell
CC cultures, by forming an Ese-Esp15 complex, then binding dynamn to the
CC complex. Generally conditions that can be treated include cancer;
CC abnormal cell division or migration; viral infection; or abnormal
CC receptor signalling, tissue development or synaptic transmission
XX
SQ Sequence 5082 BP; 1363 A; 1253 C; 1391 G; 1075 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,69e-35 Length: 5082
Score: 323.00 Matches: 58
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.31% Mismatches: 0
Query Match: 98.48% Indels: 0
DB: 3 Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x AA239008 (1-5082)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTTPaGAlaLysLysAspAsnHisLeu 20
DB 2973 GTGGAAGGGCTTCAAGCGCCCTGTATCCCTGGAGAGCCAAAAAAGACAACACCTTA 3032
QY 21 AsnPhaEnuLysAsnAspValIleThrValLeuGluGlnGlnAspMetTyrTTPheGly 40
DB 3033 AATTTTAAACAAGAGACGTATCAACGTTCTGGAACAGCAAGACATGTGGTGGT 3092
```

QY 41 GluValGlnGlyGlnIysGlyTTPheProIysSerTyrValIlystLeuIleSerGly 59
 DB 3093 GAAGTTCAGAGTCAGAAAGGCTTGTTCCCAAGCTTACGTGAACCTCATTTCAAGG 3149

RESULT 8
 ID AAZ39025 standard; cDNA; 5144 BP.
 AC AAZ39025;
 XX
 DT 28-FEB-2000 (first entry)
 XX
 DE Mouse Ese1L coding sequence.
 XX
 KM Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;
 KM regulation; actin cytoskeleton; detection; cancer; infection;
 KM EH-domain and SH3-domain regulator of endocytosis; anticancer;
 KM antiproliferative; antiviral; ss.
 XX
 OS Mus sp.
 XX
 PN MO9955728-A2.
 PD 04-NOV-1999.
 XX
 PF 27-APR-1999; 99WO-CA000375.
 XX
 PR 27-APR-1998; 98CA-02230201.
 PR 05-FEB-1999; 99US-0118739P.
 XX
 PA (HSCR-) HSC RES & DEV LP.
 XX
 PI Egan SE, Wang W, Sengar A;
 XX
 DR MPI: 2000-052802/04.
 DR P-PSDB; AAY57449.
 PT New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
 PT endocytosis, used e.g. for treating cancer or preventing viral infection.
 XX
 PS Claim 6; Page 59-62; 99pp; English.

XX The present invention specifically describes mammalian Ese1 and 2
 CC proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
 CC regulator of endocytosis). (I) are involved in regulation of clathrin-
 CC mediated endocytosis (as a complex with Esp1s protein), vesicular
 CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
 CC mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
 CC sequences antisense to the (I) polynucleotide; agents that downregulate
 CC expression of Ese genes or antagonists of an Ese binding partner are used
 CC to treat diseases associated with undesirable endocytosis and resulting
 CC changes in cellular function. Particularly overexpression of Ese1 is used
 CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
 CC administration of (I) is used to promote endocytosis of selected cells.
 CC (Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
 CC cells that can be stimulated to proliferate by a growth factor receptor;
 CC and similar compounds (also inactive Ese mutants) can be used to prevent
 CC viral infection. Endocytosis may also be regulated, in vivo or in cell
 CC cultures, by forming an Ese-Esp1s complex, then binding dynamn to the
 CC complex. Generally conditions that can be treated include cancer;
 CC abnormal cell division or migration; viral infection; or abnormal
 CC receptor signalling, tissue development or synaptic transmission. The
 CC present sequence represents mouse Ese1L coding sequence

CC Sequence 5144 BP; 1447 A; 1298 C; 1415 G; 984 T; 0 U; 0 Other;
 XX

Alignment Scores:
 Pred. No.: 1,71e-35
 Score: 323.00
 Percent Similarity: 100.00%
 Best Local Similarity: 98.31%
 Query Match: 98.48%
 DB: 3

Length: 5144
 Matches: 58
 Conservative: 1
 Mismatches: 0
 Indels: 0
 Gaps: 0

US-09-720-934-2_copy_908_966 (1-59) x AAZ39025 (1-5144)

QY 1 ValGlnGlyLeuGlnIleGlnIleValIleuTyrProTTrArgAlaIlystIysAspAenHisLeu 20
 DB 2716 GTGGAAGGGCTAACAGCGCAAGCCCTGTATCCCTGGAGAGCCAAAAAGACCAACCACTTA 2775

QY 21 AsnPhaAsnIysAsnAspValIleThrValIleuGlnGlnIleAspWecTyrTTPheGly 40
 DB 2776 AATTTTACAAAGTAGCGTCATCCGTTCTGGAACAGCAAGACATGTGTGTTTGA 2835

QY 41 GluValGlnGlyGlnIysGlyTTPheProIysSerTyrValIlystLeuIleSerGly 59
 DB 2836 GAAGTTCAGAGTCAGAAAGGCTTGTTCCCAAGCTTACGTGAACCTCATTTCAAGG 2892

RESULT 9
 ID AAZ39024 standard; cDNA; 5738 BP.
 AC AAZ39024;
 XX
 DT 28-FEB-2000 (first entry)
 XX
 DE Mouse Ese1L cDNA sequence.
 XX
 KM Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;
 KM regulation; actin cytoskeleton; detection; cancer; infection;
 KM EH-domain and SH3-domain regulator of endocytosis; anticancer;
 KM antiproliferative; antiviral; ss.
 XX
 OS Mus sp.
 XX
 PN MO9955728-A2.
 PD 04-NOV-1999.
 XX
 PF 27-APR-1999; 99WO-CA000375.
 XX
 PR 27-APR-1998; 98CA-02230201.
 PR 05-FEB-1999; 99US-0118739P.
 XX
 PA (HSCR-) HSC RES & DEV LP.
 XX
 PI Egan SE, Wang W, Sengar A;
 XX
 DR MPI: 2000-052802/04.
 DR P-PSDB; AAY57449.
 PT New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of
 PT endocytosis, used e.g. for treating cancer or preventing viral infection.
 XX
 PS Claim 6; Page 56-59; 99pp; English.

XX The present invention specifically describes mammalian Ese1 and 2
 CC proteins (I) and their splice variants (Ese = EH-domain and SH3-domain
 CC regulator of endocytosis). (I) are involved in regulation of clathrin-
 CC mediated endocytosis (as a complex with Esp1s protein), vesicular
 CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,
 CC mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);
 CC sequences antisense to the (I) polynucleotide; agents that downregulate
 CC expression of Ese genes or antagonists of an Ese binding partner are used
 CC to treat diseases associated with undesirable endocytosis and resulting
 CC changes in cellular function. Particularly overexpression of Ese1 is used
 CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while
 CC administration of (I) is used to promote endocytosis of selected cells.
 CC (Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of
 CC cells that can be stimulated to proliferate by a growth factor receptor;
 CC and similar compounds (also inactive Ese mutants) can be used to prevent
 CC viral infection. Endocytosis may also be regulated, in vivo or in cell
 CC cultures, by forming an Ese-Esp1s complex, then binding dynamn to the
 CC complex. Generally conditions that can be treated include cancer;
 CC abnormal cell division or migration; viral infection; or abnormal
 CC receptor signalling, tissue development or synaptic transmission. The

CC present sequence represents mouse Esell cDNA sequence
XX Sequence 5738 BP; 1556 A; 1456 C; 1613 G; 1113 T; 0 U; 0 Other;
SQ Alignment Scores:
Pred. No.: 1.97e-35 Length: 5738
Score: 323.00 Matches: 58
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.31% Mismatches: 0
Query Match: 98.48% Indels: 0
DB: Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x AA239024 (1-5738)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTYrProTyrArgAlaLysLysAspAsnHisLeu 20
Db 2975 GTGGAGGGGCTCAAGCTCAAGCCCTGTATCCCTGGAGACCCAAAAAGCAACCACTTA 3034
QY 21 AsnPhenylAspAspValIleThrValLeuGlnGlnAspMetTyrTyrPheGly 40
Db 3035 AATTTAAACAAAGTGAACGTCATCACCGTCTGGAAACGACACATGTGTGTTTGA 3094
QY 41 GluValGlnGlyGlnLysGlyTyrPheProLysSeryrValLysLeuIleSergly 59
Db 3095 GAAGTTCAAGGTCAAGAGGGTGTGTTCCCAAGCTTACGTGAACCTCATTTCAAGG 3151
RESULT 10
AAS84763
ID AAS84763 standard; cDNA; 7435 BP.
XX AAS84763;
AC
XX 13-FEB-2002 (first entry)
DT
XX
DE DNA encoding novel human diagnostic protein #20567.
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS Homo sapiens.
XX
XX WO200175067-A2.
PN
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
PF
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
PI Drmanac RT, Liu C, Tang YT;
DR WPI: 2001-639362/73.
DR P-PSDB; ABG20576.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1: SEQ ID NO 20567; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (II) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting (II) and as a food
CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 7435 BP; 2190 A; 1690 C; 1829 G; 1726 T; 0 U; 0 Other;
XX
XX Alignment Scores:
Pred. No.: 3.87e-34 Length: 7435
Score: 315.00 Matches: 59
Percent Similarity: 98.33% Conservative: 0
Best Local Similarity: 98.33% Mismatches: 0
Query Match: 96.04% Indels: 1
DB: Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x AAS84763 (1-7435)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTYrProTyrArgAlaLysLysAspAsnHisLe 20
Db 3019 GTGGAGGGGCTCAAGCTCAAGCCCTATATCTTGGAGAGCCAAAAAGCAACCACTT 3078
QY 20 uAsnPhenylAspAspValIleThrValLeuGlnGlnAspMetTyrTyrPheGly 40
Db 3079 AATTTTAAACAAAGATGTCATCACCGTCTGGAAACGACACATGTGTGTTTGG 3138
QY 40 yGluValGlnGlyGlnLysGlyTyrPheProLysSeryrValLysLeuIleSergly 59
Db 3139 AGAAGTTCAAGGTCAAGAGGGTGTGTTCCCAAGCTTACGTGAACCTCATTTCAAGG 3196
RESULT 11
AAS84759
ID AAS84759 standard; cDNA; 395 BP.
XX AAS84759;
AC
XX 13-FEB-2002 (first entry)
DT
XX
DE DNA encoding novel human diagnostic protein #20563.
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS Homo sapiens.
XX
XX WO200175067-A2.
PN
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
PF
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
PI Drmanac RT, Liu C, Tang YT;
DR WPI: 2001-639362/73.
DR P-PSDB; ABG20572.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX

PS Claim 1; SEQ ID NO 20563; 103bp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS84197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 395 BP; 114 A; 96 C; 100 G; 85 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8,91e-31	Length:	395
Score:	280.00	Matches:	49
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	85.37%	Indels:	0
DB:	5	Gaps:	0

US-09-720-934-2_COPY_908_966 (1-59) X AAS84759 (1-395)

QY 11 ProTTPATGAlaLysIysAspAAsnHisLeuVal 30

Db 1 CCTTGGAGAGCCAAAAGACACACCTTAATTTAACAAATGATGTCACCGTC 60

QY 31 LeuGluGlnGlnAspMetTTPPhPheGlyGluValGlnGlnIysGlyTTPPhPhePro 50

Db 61 CTGAAACGACAGACATGTGTGTGAGAGAGTTCAAGGTCAAGAGGTTGTTCCCC 120

QY 51 LysSerTyrValIysLeuIleSerGly 59

Db 121 AAGTCTTACGTGAACATTCATTCAAGG 147

RESULT 12

ABA66139

ID ABA66139 standard; DNA; 206 BP.

XX ABA66139;

AC ABA66139;

XX 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #14444.

XX Human foetal liver; gene expression; single exon nucleic acid probe; ss.

OS Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US000669.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632365.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR,

XX WPI; 2001-483447/52.

DR Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human foetal liver.

XX Claim 4; SEQ ID NO 14444; 639bp + Sequence Listing; English.

XX The invention relates to a single exon nucleic acid probe for measuring

CC human gene expression in a sample derived from human foetal liver. The

CC single exon nucleic acid probes may be used for predicting, measuring and

CC displaying gene expression in samples derived from human foetal liver. The

CC present sequence is a single exon nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 206 BP; 69 A; 33 C; 55 G; 49 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5,48e-27	Length:	206
Score:	251.00	Matches:	44
Percent Similarity:	86.44%	Conservative:	7
Best Local Similarity:	74.58%	Mismatches:	8
Query Match:	76.52%	Indels:	0
DB:	4	Gaps:	0

US-09-720-934-2_COPY_908_966 (1-59) X ABA66139 (1-206)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTTPProTTPATGAlaLysIysAspAAsnHisLeu 20

Db 10 GTAGAAAACCTTAAAGACACAGGCCCTTGTCTGTGACTGCAGAAAGATTAACCATTTG 69

QY 21 AsnPheAsnIysAspAspValIleThrValLeuGlnGlnAspMetTTPPhPheGly 40

Db 70 AACTCTCAAAACATGACATATTTACTCTTGGAGCGACAAAGAAATGTGGTGTGG 129

QY 41 GluValGlnGlnIysGlyTTPPhPheProLysSerTyrValIysLeuIleSerGly 59

Db 130 GAGGTGATGAGAGAGAGAGATGTTCCCAATCTTATGTCAAGATCATTCCTGG 186

RESULT 13

AA146323

ID AA146323 standard; DNA; 206 BP.

XX AA146323;

AC AA146323;

XX 17-OCT-2001 (first entry)

DE Probe #15009 used to measure gene expression in human placenta sample.

XX Probe; microarray; human; placenta; antenatal diagnosis;

XX genetic disorder; ss.

OS Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US000663.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632365.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 15009; 654bp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 206 BP; 69 A; 33 C; 55 G; 49 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5,48e-27 Length: 206
Score: 251.00 Matches: 44
Percent Similarity: 86.44% Conservative: 7
Best Local Similarity: 74.58% Mismatches: 8
Query Match: 76.52% Indels: 0
DB: 4 Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x AAI6323 (1-206)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTPrpArgAlaIleYlsAspAsnHisLeu 20
Db 10 GTAGAAACTTAAAGCAGAGCCCTTTCTTCTGACTGCAAGAAAGATTAACCACTTG 69
QY 21 AsnPheAsnYAsnAspValIleThrValLeuGlnGlnAspMetTrpTrpPheGly 40
Db 70 AACTCTCAAAACATGACATATTACTGCTCTGGAGCAGCAAGAAATTTGGTGGG 129
QY 41 GluValGlnGlyGlnIleGlyTyrPheProIysSerTyrValIleSergly 59
Db 130 GAGGTGCAATGAGAGAGAGATGTTTCCCAATCTTATGTCAGATCATCTCTGGG 186
RESULT 14
ABS39875
ID ABS39875 standard; DNA; 206 BP.
XX
AC ABS39875;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID No 14865.
XX
KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US000664.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 4; SEQ ID NO 14865; 658bp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC specification for this patent does not appear in the printed
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 206 BP; 69 A; 33 C; 55 G; 49 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5,48e-27 Length: 206
Score: 251.00 Matches: 44
Percent Similarity: 86.44% Conservative: 7
Best Local Similarity: 74.58% Mismatches: 8
Query Match: 76.52% Indels: 0
DB: 4 Gaps: 0
US-09-720-934-2_COPY_908_966 (1-59) x ABS39875 (1-206)
QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTPrpArgAlaIleYlsAspAsnHisLeu 20
Db 10 GTAGAAACTTAAAGCAGAGCCCTTTCTTCTGACTGCAAGAAAGATTAACCACTTG 69
QY 21 AsnPheAsnYAsnAspValIleThrValLeuGlnGlnAspMetTrpTrpPheGly 40
Db 70 AACTCTCAAAACATGACATATTACTGCTCTGGAGCAGCAAGAAATTTGGTGGG 129
QY 41 GluValGlnGlyGlnIleGlyTyrPheProIysSerTyrValIleSergly 59
Db 130 GAGGTGCAATGAGAGAGAGATGTTTCCCAATCTTATGTCAGATCATCTCTGGG 186
RESULT 15
ABS14329
ID ABS14329 standard; DNA; 206 BP.
XX
AC ABS14329;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon ORF from lung SEQ ID No 14320.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosterosi;
KW pulmonary histiocytosis; lymphangiomyomatosis; Karsageners syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX

PN WO200186003-A2.
 XX 15-NOV-2001.
 PD 30-JAN-2001; 2001WO-US0000665.
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-0063236P.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples.
 XX
 PS Claim 4; SEQ ID NO 14320; 634bp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC; the novel set of probes which hybridise at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned in the specification, or encoded by the
 CC of 12011 sequences, mentioned in the specification, a peptide comprising one
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karagenen syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a single exon probe open reading frame of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 206 BP; 69 A; 33 C; 55 G; 49 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.48e-27
 Score: 251.00
 Percent Similarity: 86.44%
 Best Local Similarity: 74.58%
 Query Match: 76.52%
 DB: 6

Length: 206
 Matches: 44
 Conservative: 7
 Mismatches: 8
 Indels: 0
 Gaps: 0

US-09-720-934-2_COPY_908_966 (1-59) x ABS14329 (1-206)

QY 1 ValGluGlyLeuGlnAlaGlnAlaLeuTyrProTTPARGAlaLysLysAspAsnHisLeu 20
 Db 10 GTAGAAACCTTAAAGACAGGCCCTTGTTCCTGAGCTGCAAGAAATACCACTTG 69
 QY 21 AsnPeAsnLysAsnAspValIleThrValLeuGluGlnGlnAspMetTTPpPheGly 40
 Db 70 AACTTCTCAAAACATGACATTTACTGTCTTGAGACAGCAAGAAATGGTGTTGGG 129
 QY 41 GluValGlnGlyGlnLysGlyTTPpPheProLysSerTyrValLysLeuIleSerGly 59
 Db 130 GAGGTGCATGAGAGAGAGATGGTTCCCAATCTTATGTCAAGATCATTCCTGGG 186

Search completed: July 1, 2004, 13:31:56
 Job time : 191.112 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 1, 2004, 12:16:50 ; Search time 1391.43 Seconds
(without alignments)
1900.151 Million cell updates/sec

Title: US-09-720-934-2_COPY_740_800
Perfect score: 333

Sequence: 1 VKVYVYRALYPFESRSHDEI.....GGEIKGKTGWFPANVAKIP 61

Scoring table:

BLAST62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-O=cmr2.1/USPTO.spool_p/US09720934/runat_30062004_064539_13430/app_query.fasta_1.1386
-DB=GenEmbl -GFWT=1step -SUFFIX=5ge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNIT3-bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcio -NORM=ext -HEASize=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720934 @CCN 1.1 7509 @runat_30062004_064539_13430 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONELOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_htg:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sts:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vt:.*
15: em_da:.*
16: em_fun:.*
17: em_hum:.*
18: em_in:.*
19: em_mu:.*
20: em_om:.*
21: em_or:.*
22: em_ov:.*
23: em_pat:.*
24: em_ph:.*
25: em_pl:.*
26: em_ro:.*
27: em_sts:.*
28: em_un:.*

29: em_vt:.*
30: em_htg_hum:.*
31: em_htg_inv:.*
32: em_htg_other:.*
33: em_htg_mus:.*
34: em_htg_pln:.*
35: em_htg_rnd:.*
36: em_htg_mam:.*
37: em_htg_vrt:.*
38: em_sy:.*
39: em_htg_hum:.*
40: em_htg_mus:.*
41: em_htg_other:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	333	100.0	306	AY254341	AY254341 Homo sapi
2	333	100.0	3231	BD205037	BD205037 Isolated
3	333	100.0	3241	HSU61166	U61166 Human SH3 d
4	333	100.0	3723	AF132478	AF132478 Mus muscu
5	333	100.0	5145	AF132481	AF132481 Mus muscu
6	333	100.0	5195	BD205035	BD205035 Isolated
7	333	100.0	5199	BD205033	BD205033 Isolated
8	324	97.3	4321	HSW86384	BX581275 Homo sapi
9	320.5	96.2	475	AY127576	AY127576 Mus muscu
10	320.5	96.2	5287	AF064243	AF064243 Homo sapi
11	320.5	96.2	5381	AF114488	AF114488 Homo sapi
12	320.5	96.2	5458	BD205034	BD205034 Isolated
13	320.5	96.2	6439	AF114487	AF114487 Homo sapi
14	320.5	96.2	7247	AF064244	AF064244 Homo sapi
15	317	95.2	4103	AF032118	AF032118 Xenopus l
16	309.5	92.9	3812	AF132672	AF132672 Rattus no
17	309.5	92.9	4025	AF127798	AF127798 Rattus no
18	216	64.9	3594	AF132480	AF132480 Mus muscu
19	216	64.9	4447	AX428899	AX428899 Sequence
20	216	64.9	4557	AF182199	AF182199 Homo sapi
21	216	64.9	4977	AF132479	AF132479 Mus muscu
22	216	64.9	5804	AK122480	AK122480 Mus muscu
23	216	64.9	5828	BD167848	BD167848 Method fo
24	216	64.9	5828	AF248540	AF248540 Homo sapi
25	216	64.9	5938	AB033082	AB033082 Homo sapi
26	216	64.9	6091	AF182198	AF182198 Homo sapi
27	215	64.6	2873	AR175271	AR175271 Sequence
28	186	55.9	169774	AC138787	AC138787 Sus scrofa
29	186	55.9	175319	AC134837	AC134837 Mus muscu
30	186	55.9	181782	AC116970	AC116970 Rattus no
31	186	55.9	217131	AC126053	AC126053 Rattus no
32	186	55.9	278501	AC098552	AC098552 Rattus no
33	186	55.9	279148	AC123507	AC123507 Rattus no
34	184	55.3	2408	AY069517	AY069517 Drosophila
35	184	55.3	3750	AF054612	AF054612 Drosophila
36	184	55.3	3999	AF053957	AF053957 Drosophila
37	184	55.3	59264	AC018278	AC018278 Drosophila
38	184	55.3	127035	AC006574	AC006574 Drosophila
39	184	55.3	160282	AC009253	AC009253 Drosophila
40	184	55.3	273995	AE003669	AE003669 Drosophila
41	183	55.0	142974	AC129075	AC129075 Felis cat
42	183	55.0	178801	AC121075	AC121075 Canis fam
43	183	55.0	184526	AC144367	AC144367 Papio ann
44	183	55.0	194157	AC144368	AC144368 Gallus ga
45	183	55.0	203423	AC137544	AC137544 Gallus ga

RESULT 1

ALIGNMENTS

AY254341
LOCUS Homo sapiens intersec1 306 bp mRNA linear PRI 22-APR-2003
DEFINITION 1 isoform 5 (TTSN1) mRNA, partial cds,
alternatively spliced.
ACCESSION AY254341
VERSION AY254341.1 GI:30060229
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Teyba, L., Skrypkina, I., Nikolaenko, A., Fortna, A., Rynditch, A. and
Gardiner, K.
TITLE Tissue-specific regulation of intersec1 1 transcriptional
isoforms in human and mouse
JOURNAL
REFERENCE 2 (bases 1 to 306)
AUTHORS Teyba, L., Skrypkina, I., Nikolaenko, A., Fortna, A., Rynditch, A. and
Gardiner, K.
TITLE Direct Submision
JOURNAL
COMMENT Submitted (12-MAR-2003) Department of Molecular Oncogenetics,
Institute of Molecular Biology and Genetics of National Academy of
Sciences of Ukraine, 150 Zabolotnogo Str., Kyiv 03143, Ukraine
FEATURES
source
1. .306
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
/tissue_type="kidney"
/dev_stage="12 week-old embryo"
<1. .>306
/gene="TTSN1"
<1. .>306
/gene="TTSN1"
/note="alternatively spliced"
/codon_start=1
/product="intersec1 1 isoform 5"
/protein_id="AAP13099.1"
/db_xref="GI:30060230"
/translation="QCEPAPKPAVQAPWSTAEKGPLRTISAQENVKVVYRYALPPEER
SHDRTITQPGDIWVWDESQTGEPGMLGSELKGTGMPANAEKIPENEVAPVKEPTV"
ORIGIN
Alignment Scores: 3.01e-34 Length: 306
Pred. No.: 333.00 Matches: 61
Score: 333.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 9 Gaps: 0
US-09-720-934-2_COPY_740_800 (1-61) x AY254341 (1-306)
QY 1 Vallysvalvaltyrtyrarghaleuuyrprophglsuserarqserhisaspgluile 20
DB 88 GTHAAAGGTGTATTAACCGGCACTGTACCCCTTGATCCAGAACCATGATGTAATC 147
QY 21 ThrllleglnproglYAspilleValmetValaspgluserGlnThrglYglupProglYTP 40
DB 148 ACTATTCACGCCAGACAGACATAGTCATGATGATGAAGCCAGACTGAGAACCCGCTGG 207
QY 41 LeuGlYglYgluLeuYsglYysThrglYTPpPheProAlAAsnTYzAlagluYsile 60
DB 208 CTTGGAGGAGAAATTAAGAAAGACAGAGGCTGTCTCCGCAAACTATGCAAGGAAATC 267
QY 61 Pro 61
DB 268 CCA 270

RESULT 2
BD205037 3231 bp DNA linear PAT 17-JUL-2003
LOCUS Isolated SH3 gene relating to myeloproliferative disorders and
DEFINITION leukemia and utilization thereof.
ACCESSION BD205037
VERSION BD205037.1 GI:33014807
KEYWORDS JP 2002511267-A/5.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Korenberg, J.R. and Chen, X.N.
TITLE Isolated SH3 gene relating to myeloproliferative disorders and
leukemia and utilization thereof
JOURNAL
COMMENT Patent: JP 2002511267-A 5 16-APR-2002;
CEDARS SINAI HEALTH SYSTEM ET AL
OS Homo sapiens (human)
PN JP 2002511267-A/5
PD 16-APR-2002
PF 16-APR-1999 JP 2000543610
PR 16-APR-1998 US 60/082007
PI JULIE R KORENBERG, XIAO NING CHEN
PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10,
PC C1201/68,
PC G01N33/68//A61K48/00, C12N15/00, C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and
leukemia
CC and utilization thereof.
FH Key Location/Qualifiers
FT source 1. .3231
/organism="Homo sapiens (human)"
Location/Qualifiers
1. .3231
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 3.75e-33 Length: 3231
Pred. No.: 333.00 Matches: 61
Score: 333.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 6 Gaps: 0
US-09-720-934-2_COPY_740_800 (1-61) x BD205037 (1-3231)
QY 1 Vallysvalvaltyrtyrarghaleuuyrprophglsuserarqserhisaspgluile 20
DB 710 GTHAAAGGTGTATTAACCGGCACTGTACCCCTTGATCCAGAACCATGATGTAATC 769
QY 21 ThrllleglnproglYAspilleValmetValaspgluserGlnThrglYglupProglYTP 40
DB 770 ACTATTCACGCCAGACAGACATAGTCATGATGATGAAGCCAAACTGAGAACCCGCTGG 829
QY 41 LeuGlYglYgluLeuYsglYysThrglYTPpPheProAlAAsnTYzAlagluYsile 60
DB 830 CTTGGAGGAGAAATTAAGAAAGACAGAGGCTGTCTCCGCAAACTATGCAAGGAAATC 889
QY 61 Pro 61
DB 890 CCA 892
RESULT 3
HSU61166 3241 bp mRNA linear PRI 23-JUL-1996
LOCUS Human SH3 domain-containing protein SH3P17 mRNA, complete cds.
DEFINITION
ACCESSION U61166
VERSION U61166.1 GI:1438932

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Sparke,A.B., Hoffman,N.G., McConnell,S.J., Fowlkes,D.M. and Kay,B.K.
TITLE Cloning of ligand targets: systematic isolation of SH3 domain-containing proteins
JOURNAL Nat. Biotechnol. 14 (6), 741-744 (1996)
MEDLINE 98294438
PUBMED 9630982
REFERENCE 2 (bases 1 to 3241)
AUTHORS Pirozzi,G., McConnell,S.J., Uveges,A. and Fowlkes,D.M.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-1996) CYTOGEN Corp., 307 College Road East, Princeton, NJ 08540, USA
FEATURES
source Location/Qualifiers
1. 3241
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="bone marrow"
37. 1599
/codon_start=1
/product="SH3 domain-containing protein SHP17"
/protein_id="AAC50592.1"
/db_xref="GI:1438933"
/translation="MEAEERLKQKEQERKIIIELEKQEAQRAQERDKOMLTHVOQED
EHORPKLHEEELKREESVKKKQDEKQKQKQADLGRLEFHQEPAPVQAPST
AEKGLPTISAQENVKVVYRALYPPESRSHDEITIOGDIVWDESGTEPGLGSEL
KQGTGFPANVAKIKPENVPAPKVPVSTSTAPAKLALRETPALATYSSSPSTP
NNWADSSITPSTSTNEKPEKTDMDMAAQAQPSAGOLRQASATPATATGSSSP
VGGGKRVGLQAQALYPAKADNNKNDVITVLEQDMMWFGEVQGGQGWPEK
SYVKLISGPIRSTSDSSSPSLKRVASPAKPVVSGEIAQIVASVATGPEQ
LTLAEQQLILIRKKNPGGWEELQARGRQIQGMPANVYKLSPTSKITPTTEPK
STALAAVCQVIGMTDYTAQNDELAFNKGQIINVINKEDPDMWKGVEVNGVGLFPSNY
VLTMDPSQ"

ORIGIN
Alignment Scores:
Pred. No.: 3.76e-33 Length: 3241
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-720-934-2_COPY_740_800 (1-61) x HSU61166 (1-3241)

QY 1 VallysValValTYrTYrAArgAlaLeuTYrProPheGluSerArgSerHisAspGluIle 20
Db 382 GTAAAAAGTGCTATTATCCGGGCACTGTAACCCCTTGATCCAGAGCCATATGAAATC 441
QY 21 ThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr 40
Db 442 ACTATCCAGCCAGAGACTAGTCAATGATGATGAAGCAAACTGCGAACCCTGG 501
QY 41 LeuGlyGlyGluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIle 60
Db 502 CTGGAGGAGATTAAGAAAGAAAGACAGGCTGCTCCCTGCAAACTATGACAGAAATC 561
QY 61 Pro 61
Db 562 CCA 564

RESULT 4
AF132478 3723 bp mRNA linear ROD 09-MAR-1999
LOCUS Mus musculus Esel protein mRNA, complete cds.
DEFINITION AF132478
ACCESSION AF132478
VERSION AF132478.1 GI:4378884

KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scurionahli; Muridae; Murinae; Mus.
AUTHORS Sengar,A.S., Wang,W., Bishay,J., Cohen,S. and Egan,S.E.
TITLE The EH and SH3 domain Ese proteins regulate endocytosis by linking to dynamin and Eps15
JOURNAL EMBO J. 18 (5), 1159-1171 (1999)
MEDLINE 99164083
PUBMED 10064583
REFERENCE 2 (bases 1 to 3723)
AUTHORS Sengar,A.S., Wang,W., Cohen,S., Bishay,J. and Egan,S.E.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-1999) Programs in Cancer & Blood Research/Developmental Biology, The Hospital for Sick Children, 555 University Avenue, Toronto, ON M5G-1X8, Canada
FEATURES
source Location/Qualifiers
1. 3723
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1. 3542
/codon_start=1
/product="Esel protein"
/protein_id="A019746.1"
/db_xref="GI:4378885"
/translation="MAQFPPTPGSGIDWVAITYEERAKHDQQLSLKPIAGRTTGDA
RNFFRSGHGPQVLAQTNALADNNDRNDQVFFSTAMKILIKLIGYQLPSTLPVM
KQPPAIVSSAFAGIGIASMPLEIAVAAPWMSIIVGCGSPVLSVPAAPAAVPLAN
GAPVPIQPLPAFAHPAPATPKSSFSRSGSGLNKLQKAGSFVSAVPAAPAAVPLAN
QSRRLKQLFNSHDKTMSGLTGPQARTILMOSLPOQLAIVNLSIDIDQGLTLA
BEFILAMHLIDVAMSGQPLPVPILPEPIPTSPRRVRSVSGMSVYSSSSQDRIPEPS
SEDOOEKKLVTTFEDKRENNPEEGSVLEKRRQALLDQQRQRLQLEAEQER
KREBOEQAKQLELEKQLEKQRELEERREKREIRERREKLELRQLEQER
NRQELNQRKQEGTVLKARRTLELELADNKKQLBGLQDINCRLATQRE
LESTNKSRELRLAEITTHLQOQLQESQMLGRIIPKQILSDQLKQVQNSLHSDSLT
IKRALBAKELARQQLREQLDEVERETRSKLOIDVNNQLEKLEIHSQQLQOKQSL
EAVRLKQKEERKSELEKQEDQARVQERDKOMLEHVOEQEOPRPRPHEDRLKR
EDSVRKKEERAKPEMDKQRLHHPHDEPAKLATQAWSTTEKGLPTTISAESVY
VYVRYALYPPESRSHDEITIOGDIVWDESGTEPGLGELKGTGMPANVAKIK
KEEVTTPAKVDLTLSAPAKLALRETPALYPPESRSHDEITIOGDIVWDESGTEP
KRETNWMTWAAQPSLTVPAGQLRQSAFTATATGATGATGATGATGATGATGATG
LYPAKKNHNLNFKNSDVITVLEQDMMWFGEVQGGQGWPEVYKLSPTSKITPTTEPK
DTGPESPASLKRVASPAKPAIPGEPIFLAMTYTSSSGGLTPOCGDITVYTKGCD
MWITGVKSGVPSNYVIRKDSGSGTAGTSGISGKREIQAIVASTATGPEQQLT
APGQLILIRKKNPGGWEELQARGRQIQGMPANVYKLSPTSKITPTTEPKTAV
QPAVCQVIGMTDYTAQNDELAFNKGQIINVINKEDPDMWKGVEVNGVGLFPSNYKL
TTMDPSQ"

ORIGIN
Alignment Scores:
Pred. No.: 4.36e-33 Length: 3723
Score: 333.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-720-934-2_COPY_740_800 (1-61) x AF132478 (1-3723)

QY 1 VallysValValTYrTYrAArgAlaLeuTYrProPheGluSerArgSerHisAspGluIle 20
Db 2212 GTAAAAAGTGCTATTATCCAGGCGCTGTAACCCCTTGATCCAGAGCCATATGATG 2271
QY 21 ThrIleGlnProGlyAspIleValMetValAspGluSerGlnThrGlyGluProGlyTyr 40
Db 2272 ACCATCCAGCCAGAGATATAGTCAATGATGAAGCACTGCGAGACCCAGAGATGG 2331
QY 41 LeuGlyGlyGluLeuLysGlyLysThrGlyTyrPheProAlaAsnTyrAlaGluLysIle 60
Db 2332 CTGGAGGAGAGCTGAAAGGAAAGACGAGATGTTCCCTGCAAACTATGACGAAATG 2391

polYA_site 4289
/gene="DKFZp686d17173"

ORIGIN

Alignment Scores:
Pred. No.: 7.55e-32 Length: 4321
Score: 324.00 Matches: 60
Percent Similarity: 98.36% Conservative: 0
Best Local Similarity: 98.36% Mismatches: 1
Query Match: 97.30% Indels: 0
Gaps: 0
DB: 9

US-09-720-934-2_COPY_740_800 (1-61) x HSM806384 (1-4321)

QY 1 VallyeValValYrYrYrAgaAlaLeuTyProPheGluSerArgSerHisaaPgluIle 20
1785 GTAAAGTGTATATTACCGGCACTGTACCCCTTGAATCCAGAAAGCCATGATGAATC 1844

Db 21 ThrTleGlnProGlyaPllleValMetValaaspGluSerGlnThrGlyGluProGlyTyr 40
1845 ACTATCCAGCCAGAGACATAGTCATGTGATGAAGCTAAAGCTGGAAGAACCCGCTGG 1904

QY 41 LeuGlyGluGluLeuLysGlyLysThrGlyTyrPheProAlaSerYrAlaGluLysIle 60
1905 CTGGAGAGAAATTAAAGAAAGACAGGGGTGTTCCCTGCAACTATGACAGAAAAATC 1964

Db 61 Pro 61
1965 CCA 1967

QY 61 Pro 61

Db 1965 CCA 1967

RESULT 9
AY127576 475 bp mRNA linear ROD 01-DEC-2002
LOCUS Mus musculus intersec tin isoform 5 (Itsn) mRNA, partial cds;
DEFINITION alternatively spliced.
ACCESSION AY127576
VERSION AY127576.1 GI:25989570
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 475)
JOURNAL Slavov, D., Rynditch, A. and Gardiner, K.
AUTHORS Skripkina, I., Tsyba, L.,
TITLE Novel splicing forms of human and mouse intersec tin gene
transcripts
JOURNAL Ukr. Biokhim. Zh. (2002) In press
REFERENCE 2 (bases 1 to 475) In press
AUTHORS Skripkina, I., Tsyba, L. and Rynditch, A.
TITLE Direct Submission
JOURNAL Department of Molecular Oncogenetics,
Submitted (27-JUN-2002) Institute of Molecular Biology and Genetics of National Academy of
Sciences of Ukraine, 150, Zabolotnogo Str., Kiev 01143, Ukraine
FEATURES
source Location/Qualifiers
1..475
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/cx57BLP1"
/db_xref="taxon:10090"
/chromosome="14"
/tissue_type="brain"
/tissue="brain"
/gene="Itsn"
/gene="Itsn"
/note="alternatively spliced"
/codon_start=1
/product="intersec tin isoform 5"
/protein_id="AA02285.1"
/db_xref="GI:25989571"
/translation="MHQEPAKLATQAPWSTTEKGLPLTISAQSEVAVVYRLLYPPES
RSHRTTIOGPDIWYKGGWVBSQSGEFGWIGELKGTGMPFANYPANAKIPENVEPT
PAKPTDLSAPAKLATRETRAPLPVTSSEPTTPNNMADSSTWSSSNEKPEP"

ORIGIN

Alignment Scores:
Pred. No.: 2.02e-32 Length: 475
Score: 320.50 Matches: 61
Percent Similarity: 92.42% Conservative: 0
Best Local Similarity: 92.42% Mismatches: 0
Query Match: 96.25% Indels: 5
DB: 10 Gaps: 1

US-09-720-934-2_COPY_740_800 (1-61) x AY127576 (1-475)

QY 1 VallyeValValYrYrYrAgaAlaLeuTyProPheGluSerArgSerHisaaPgluIle 20
91 GTAAAGTGTATATTACCGAGCGCTGTACCCCTTGAATCCAGAAAGCTCAGATGATC 150

Db 21 ThrTleGlnProGlyaPllleValMetValaaspGluSerGlnThrGlyGluProGlyTyr 35
151 ACCATCCAGCCAGAGATATAGTCATGTGTTAAAGGGAATGGGTGATGAAAGCCAGACT 210

QY 36 GlyGluProGlyTyrPheGlyGlyGluLeuLysGlyLysThrGlyTyrPheProAlaasn 55
211 GGAGAGCCAGATGCGCTTGGAGAGAGAGCTGAAGGAAAGACGGATGTTCCCTGCAAA 270

Db 56 TyrAlaGluLysIlePro 61
271 TATGCAAAAAGATTCCA 288

QY 56 TyrAlaGluLysIlePro 61

Db 271 TATGCAAAAAGATTCCA 288

RESULT 10
AF064243 5287 bp mRNA linear PRI 21-NOV-1998
LOCUS Homo sapiens intersec tin short form mRNA, complete cds.
DEFINITION AF064243
ACCESSION AF064243.1 GI:3859852
VERSION AF064243.1 GI:3859852
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE Homo sapiens
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 5287)
JOURNAL Guipponi, M., Scott, H.S., Chen, H., Schebesta, A., Rossier, C. and
AUTHORS Antonarakis, S.E.
TITLE Two isoforms of a human intersec tin (ITSN) protein are produced by
brain-specific alternative splicing in a stop codon
Genomics 53 (3), 369-376 (1998)
JOURNAL 99017974
AUTHORS 9799604
TITLE 2 (bases 1 to 5287)
JOURNAL Guipponi, M., Scott, H.S., Chen, H., Schebesta, A., Rossier, C. and
AUTHORS Antonarakis, S.E.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-1998) Genetics and Microbiology, CMU, 1 rue
Michel-Servet, Geneva 4 CH-1211, Switzerland
FEATURES
source Location/Qualifiers
1..5287
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
/tissue_type="brain"
/dev_stage="fetus"
/dev_stage="fetus"
/codon_start=1
/product="intersec tin short form"
/protein_id="AAC78610.1"
/db_xref="GI:3859853"
/translation="MAQPTPGGSLDIWALTVERAKDQFHSUXKPTSGPTTGDQ
RNFPQSGLPQVLAQIWAIDAMNDGMDQVFSIAMKLTKKQGYQUPPALPVM
KQOPVASAPFGMGKIASMPPLTAVAPVMSIPVGMSPPLVSSVPTAAVPIAN
GAPVYQPIPAFAHPAAATPKSSFSRSGSGSLNKLQKASFDVASVPPVEMAVP
QSRUKYRULPNSHDKTMSGHITGPARTIMQSSIPQALASINMLSDIDQGLTA
EERTIAMHLIDVAMSGDLPVLIPEYITPPSPFRVSSGSISVTSVQDRLPEEPV

Db 2666 TATGACAGAGAAATCCCA 2683
RESULT 12 4458 bp DNA linear PAT 17-JUL-2003
BD205034
LOCUS Isolated SH3 gene relating to myeloproliferative disorders and
DEFINITION leukemia and utilization thereof.
ACCESSION BD205034.1 GI:33014804
VERSION BD205034.1
KEYWORDS JP 2002511267-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 5458)
AUTHORS Korenberg, J.R. and Chen, X.N.
TITLE Isolated SH3 gene relating to myeloproliferative disorders and
leukemia and utilization thereof
JOURNAL Patent: JP 2002511267-A 2 16-APR-2002;
CEDARS SINAI HEALTH SYSTEM ET AL
COMMENT OS Homo sapiens (human)
PN JP 2002511267-A/2
PD 16-APR-2002
PF 16-APR-1999 JP 2000543610
PR 16-APR-1998 US 60/082007
PI JULIE R KORENBERG, XIAO NING CHEN
PC C12N15/09, A01K67/027, C07K14/47, C07K16/18, C07K19/00, C12N5/10,
C1201/68,
PC G01N33/68//A61K48/00 C12N15/00, C12N5/00
CC Isolated SH3 gene relating to myeloproliferative disorders and
leukemia
CC and utilization thereof.
FH Location/Qualifiers
FT source 1..5458
FT /organism="Homo sapiens (human)".
FT /db_xref="taxon:9606"
FEATURES
source 1..5458
Location/Qualifiers
1..5458
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. NO.: 2.76e-31 Length: 5458
Score: 320.50 Matches: 61
Percent Similarity: 92.42% Conservative: 0
Best Local Similarity: 92.42% Mismatches: 0
Query Match: 96.25% Indels: 5
Gaps: 1
DB: 6
US-09-720-934-2_COPY_740_800 (1-61) x BD205034 (1-5458)
QY 1 VallysvaValaYTYTYArgAlaLeuYrProphGluSerArgSerHisaspGluile 20
Db 2484 GTAAAGGGGTATTATACCGGACACTGTAACCCCTTGATTCAGAAAGCATTATGAAAC 2543
QY 21 ThrilegInProGlyVAspIleValMet-----ValaspGluSerGlnthr 35
Db 2544 ACTATCCAGCCAGGAGACATATCATGTTAAAGGGGATGGGTGGATGAAAGCCAAACT 2603
QY 36 GlyIuProGlyTYrPleuGlyGlyIuLeuLeuGlyLyTyThGlyTYrPheProAlaasn 55
Db 2604 GGAGAACCCGGCTGGCTTGAGAGAGAAATTAAGAAAGAACAGAGGCTGTTCCCTGCAAC 2663
QY 56 TYrAlaGluIuYseIlePro 61
Db 2664 TATGACAGAGAAATCCCA 2681
RESULT 13
AF114487 6439 bp mRNA linear PRI 16-JUL-2002
LOCUS AF114487

DEFINITION Homo sapiens intersecin long isoform (ITSN) mRNA, complete cds.
ACCESSION AF114487
VERSION AF114487.1 GI:4808822
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 6439)
AUTHORS Pucharcos, C., Fuentes, J.J., Casas, C., de la Luna, S., Alcantara, S.,
Arbones, M.L., Soriano, E., Estivill, X. and Pritchard, M.
TITLE Multi-splice cloning of human Intersectin (ITSN), a putative
multivalent binding protein expressed in proliferating and
differentiating neurons and overexpressed in Down syndrome
Eur. J. Hum. Genet. 7 (6), 704-712 (1999)
JOURNAL MEDLINE
PUBMED 99415290
REFERENCE 2 (bases 1 to 6439)
AUTHORS Pucharcos, C., Fuentes, J.J., Pritchard, M. and Estivill, X.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1998) Medical and Molecular Genetics Center,
Cancer Research Institute, L'Hospitalet de l'lo., Avia.
Castelldefels km. 2,7, Barcelona 08907, Spain
FEATURES
source 1..6439
Location/Qualifiers
1..6439
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
1..6439
/gene="ITSN"
269..5434
/gene="ITSN"
/codon_start=1
/product="Intersectin long isoform"
/protein_id="A4D29952.1"
/translation="MAQPTPTFGSLDITVAITVEERAKHQDSHLKPSGFTITGDQA
RNPFPGSLDPOPVLVAIWMALNNDNDMDQDVEFSIMKLTKLQGLQVLSALPPLM
KQPVATSSAPAPGMGIGIAPMPLTAAPVAVPVGMSFYTAASVPTAAVPLAN
GAPVITDLPAPAPAPANTPKSSFGSSGPGSOLNTKQKQSDGVAAPVAEAVP
QSSRLKRYQLFNHSDKTVSGHITGPQARTIMQSSLPQAGLAIWNTSIDDDGLTA
EFTIAMHLIDVAMSGQPLPPLPEVTPSFRVRSQSGISVTSSTVDQRLPREPV
LDEQQOLEKLPTVTFEDKKRNFERNMILEKRRQALLERAKREQLAQRERQZ
RKERROBERKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQLEKQLE
NNRQELNNKKEQEDIVLAKAKKTLEFLEALNDKQLEKQLEKQLEKQLEKQLE
EISTNKRRRIETIHTLOOLQDSQMLQLEIDIENNOLKEIRHNKQLOKRS
TLKRLKELARQHLRQDLBVEKRTSKLOEIDIENNOLKEIRHNKQLOKRS
MEERLKKQEKQERKITEKQERKQERKQERKQERKQERKQERKQERKQERKQERK
KREESVKKQDEKQERKQERKQERKQERKQERKQERKQERKQERKQERKQERKQERK
KVYRYALYPESRSHDEITDOPDIIWVKQHWDESGTGEWAGGELKGTGTPA
NYARKIPENEVAPKPTVDSTSAAPKLALRETPAPLAVTSSESTTPNNMADSSST
WPTSTNBERPDKNDNHLNPNKNDVITVLEQDDMMWFGQVQKQMPKSYVLLSGP
GLQDALPWRKAKNDNHLNPNKNDVITVLEQDDMMWFGQVQKQMPKSYVLLSGP
IKRSTMSQSSSESPASLKRVAAPAKVVSSEETIATYTPSSRQGLTPOGQVIL
VTXKGDWMTGTGVDGKAGVPSNVYRLQDSGSGTAGCTGKGPETLAQVASTAT
GPEOLTTLAPGOLILIRKXNPGMWGEGELQARKKRQIQMPNNYVXLSPGSKLPT
BPKSTALAAVOVIGMTDYAONDELAFNKQIQINLVNKEDPMWKGEEVGVGLF
FQPLMESBELTKXVAMT FVWKEKLINCKILKALRYRKSGKGMKPVKIGILS
AOLPMOPPIRPSQRLNGALILQCKTDEADPFKPFVRYRLAMPDCKGMPLSSPLIKP
MQRVRYRPLIKNILENTPENHPDHSILKHALERKABELSCQVNEGYREKRENDPDEMI
QAHVQCEGSEQLVNSVYNCIGPKRFLHSGLYAKSKKELVGLFNDPLLLQITLK
PLGSSGTQVPSFKSNLYQRYKPTIPINEVVLKPTPDSGDEPFIHSHIDRVYTLR
ASINVERKQKQKIAASELYETEKREKREKLYLVNSQAPATIGLWAVVAGIEADQ
CRSHGKNPYCVYTWGSOCHITKTITQDILNTPMNSNCOPTFIDLBQEVLCITVFEEDQ
FSPDDFLKRTETRVADIKDQSGSKPVTYKCLLHVEYPTGEIYVRDILQLFDEP"

EST
E
2

mis
mis
mis

mis
mis
mis
mis
mis
mis

Alignment
Pred. No
Score:
Percent S
Best Loca
Query Mah
DB:

QY
Db
QY

QY
Db
QY
Db
RESULT 15
AF032118
LOCUS
DEFINITION
ACCESSION
VERSION

